The production of syllables in connected speech

Monika Baumann

THE PRODUCTION OF SYLLABLES

IN CONNECTED SPEECH

Layout: Monika Baumann

Cover design: Bernadette Schmitt

Printed and bound by: Benda drukkers, Nijmegen

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Baumann, Monika

The production of syllables in connected speech / Monika Baumann. - [S.1 : s.n.]
(Nijmegen : Benda drukkers). - Ill.
Thesis Katholieke Universiteit Nijmegen. - With ref. - With summary in Dutch.
ISBN 90-9008946-2
Subject heading: speech production

THE PRODUCTION OF SYLLABLES

IN CONNECTED SPEECH

een wetenschappelijke proeve op het gebied van de Sociale Wetenschappen

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op maandag 5 februari 1996
des namiddags te 3.30 uur precies

door

MONIKA BAUMANN

geboren op 27 maart 1967 te Aken (Duitsland)

Promotor: Prof. dr. W.J.M. Levelt
Co-promotor: Dr. A.S. Meyer

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, Munich (Germany) and carried out at the Max-Planck-Institut für Psycholinguistik, Nijmegen (The Netherlands).

ISBN 90-9008946-2

TABLE OF CONTENTS

Acknowledgements

I've just realized that looking forward to writing the acknowledgements for your thesis in happy anticipation is much easier than actually writing them. Anyway, I owe a lot of thanks to many people who in different ways contributed to the work reported in this book.

I would like to thank Pim Levelt, my promotor, for talking to me about "strategy" at the right moments in time and for his support.

Antje Meyer intensively and excellently fed me, the theoretical linguist, with experimental psychology and statistics. As my co-promotor, she carefully read every paper I wrote and always had time to discuss things and came up with many suggestions and critical questions. Co-operating with Antje was both effective and fun.

Herbert Baumann and John Nagengast provided valuable technical support for the experiments, Christa Hausmann-Jamin kept the system running, and Ad Verbunt helped optimizing the equipment for writing and printing. Thanks also to the other people from the technical group who were always willing to help, and Richard Piepenbrock of "CELEX". Nora van Eenennaam and especially Paul Verleg called more than 800 participants - I think Paul lost some nanometers of skin dialing all those telephone numbers. Moreover, while waiting at Paul's reception desk for the participants to come (or not to come ...), I learned - between two incoming phone calls - quite a lot about Dutch language, people, history and geography over the years.

Jos van Berkum, Geert Booij, Etta Drews, T.A. Hall, Jörg Jescheniak, James McQueen, Ardi Roelofs, Bernadette Schmitt, Herbert Schriefers, Miranda van Turennout, Linda Wheeldon, and Pienie Zwitserlood helped me in thinking about many of the issues in this thesis. Cecile Kuijpers and Daan Broeder gave support with formants and transitions. Silvia Aal, Jos van Berkum, Ger Desserjer, Marc Fleischheuers, Alette Haveman, Paul Verleg, and Marianne Wissink carefully checked lists of Dutch materials for me. Agnes Bolwiender, Ingrid Schophuis, and Miranda van Turennout provided the spoken stimuli. André Linssen ran the sessions of the main experiments of Experiments 1 and 2 when he was a trainee at the Institute.

I owe many thanks to Jos van Berkum, who translated the Dutch summary that you find at the end of this book, although he was busy finishing his own thesis at the time. The jumping coda consonant on the front cover would have never come out of its syllable bubble without the graphic skills of Bernadette Schmitt.

A whole bunch of nice fellow PhD students and colleagues contributed in many ways towards making my stay at the institute a very pleasant one. Heike Behrens, the excellent cook and perfect host, provided valuable psychological and food support. Jörg Jescheniak and Janpeter de Ruiter, my subsequent officemates, heroically tolerated (and lucky enough survived) my need for fresh air.

A final "Danke!" to my family and friends back home for keeping in close contact all the time and cheering me up whenever necessary - and not only then!

die zeit vergeht

lustig
luslustigtig
lusluslustigtigtig
luslusluslustigtigtigtig
lusluslusluslustigtigtigtigtig
luslusluslusluslustigtigtigtigtigtig
lusluslusluslusluslustigtigtigtigtigtigtig
luslusluslusluslusluslustigtigtigtigtigtigtigtig

(ernst jandl)

INTRODUCTION

The syllable is a unit that phoneticians have a hard time defining, but that nevertheless almost all phonologists have incorporated into their theories as a useful concept. Psycholinguists have collected experimental evidence for the syllable. The first part of the Introduction deals with the syllable in phonetics, phonology, and psycholinguistics. It furthermore includes a description of syllables in Dutch.

The second part of the Introduction deals with the production of syllables. In particular, Levelt's model of phonological encoding will be introduced. Predictions drawn from this model were tested in the experiments reported in the remaining chapters.

The Syllable

Imagine you are participating in one of these games where the entertainment increases with the number of friends participating. It is your turn, and you are shown a little card with a concept printed on it that you have to explain without naming it, while your team has to guess what is printed on your card as quickly as possible. Imagine furthermore, that on the card that was just handed to you it said 'SYLLABLE'. The chance is high that in order to help making your team win, you will not start a report on suprasegmental phonology, sonority hierarchies of phonological segments, or stress assignment - especially since it is rather unlikely that all members of your team are phonologists. Instead you will think of some nice long words, like *communication*, *paleontology*, or *chocolate flavored biscuit*, and pronounce them carefully, with pauses between the syllables. Depending on your temperament you may accompany this by emphatic cutting gestures with you arms: *pa--le--on--to--lo--gy*. Your team should be successful at least after the second word - the syllable is an intuitively plausible unit of speech.

This thesis deals with the question of how a speaker produces the syllabic structure of an utterance. The thesis consists of two parts that address this issue from two different perspectives:

Following Levelt (1992, 1993), syllables are created at a late point during phonological encoding, while at an earlier point in time only the unsyllabified ordered sequence of a word's phonological segments is available, for instance, the segments /g, ɛ, ɪ, v/ for the word *gave*. The first part of this thesis reports a series of priming experiments investigating the time course of syllabification during word form encoding. Participants had to produce a target word as reaction to a visually presented cue. Auditory syllable stimuli were presented while the participant produced the target. The interfering stimulus either corresponded to the target's first surface syllable or it did not. The point in time when the syllable was presented was varied with respect to the visual presentation that triggered the participant's response. If syllables are produced late during phonological encoding, the syllable structure of the target word should be important when the interfering stimulus is presented at a late point in time, but not when it is presented early.

The experiments in the second part examined whether a speaker exclusively produces those syllables that surface in speech. When a sequence of several words is uttered in connected speech, the surface syllables do not necessarily coincide with word boundaries. For instance, *gave it* will syllabify as $(\text{gɛi})_\sigma(\text{vɪt})_\sigma$. One could also assume that a speaker first syllabifies the individual words, for instance, $(\text{gɛiv})_\sigma$ and $(\text{ɪt})_\sigma$. In a next step, the syllables are concatenated and the /v/ goes into the second syllable, resulting in $(\text{gɛi})_\sigma(\text{vɪt})_\sigma$. This view, for instance adopted in Lexical Phonology (Kiparsky, 1985; Kaisse & Shaw, 1985), involves two levels of syllables: The abstract syllables for the individual words, and the surface syllables. From a processing point of view, however, it seems inefficient to produce syllables that never surface, given the speed at which surface syllables are produced (three to five syllables per second). Therefore, Levelt (1992, 1993) argues that no other syllables than the surface syllables are produced during phonological encoding.

This claim was tested in production and perception experiments on syllable-final devoicing in encliticized forms. The results are accompanied by acoustic measurements on the voicing facts.

But first I will sketch the attempts to find evidence and accounts for the syllable other than its intuitive plausibility mentioned above. These come from the areas of phonetics, phonology, and psycholinguistics.

The Syllable in Phonetics

"Everyone knows that "a syllable is what *syllable* has three of" (Lass, 1984, p. 248). But despite its intuitive plausibility, there is some debate in the literature as to the syllable's phonetic definition. Or as Ladefoged (1993, p. 244) put it: "Although nearly everybody can identify syllables, almost nobody can define them". Attempts to phonetically define the syllable either look at acoustic properties of speech or at articulatory phenomena.

Articulation starts with the initiation of the air flow that is needed to produce a sound. Some organ has to produce pressure in adjacent parts of the vocal tract to make the air stream. For most sounds, the lungs will be the source of power that produces the airstream. The airstream is then modulated on its way through the articulatory apparatus by the articulatory organs, for instance the tongue or the lips, until it exits through mouth, nose, or both. To produce an [m], for instance, the lips have to be closed. Moreover, the velum has to lower such that air exits through the nose. For the central vowel schwa ([ə]), to give another example, the velum has to close the nasal cavity, the lips are open, and the tongue is in a resting position. Articulatory definitions of the syllable differ with respect to the point in the articulation process they refer to.

Definitions on the basis of sound initiation processes look at the start of articulation. The motor theory of the syllable claimed, for instance, that syllables correlate with chest pulses. A chest pulse is a contraction of the muscles of the rib cage that pushes more air out of the lungs (Stetson, 1951[2]). According to the motor theory, syllables are hence independent muscular gestures. The theory was disproved experimentally: Sometimes, one chest pulse spans two syllables, sometimes a syllable contains two chest pulses. Catford's (1992[4]) definition looked also at the early point of articulation. The definition is based on the activity of organs that produce the pressure needed to initiate the airstream - for English or Dutch this will mean the lungs. He defined syllables as minimal pulses of initiator activity that are bounded by momentary retardations. These retardations are usually caused by consonant-type articulator activity. An example for articulator activity is the movement of the tongue tip towards the area behind the upper teeth when a speaker produces a [t].

Other phoneticians tried to define the syllable by considering a later point during articulation. These accounts look at the way the airstream is modulated by the articulatory organs on its way through the vocal tract. Several articulatory organs are involved in producing a sequence of speech. Not all of them contribute in the same way to every sound, and often their activity overlaps in time. The articulation of a sound is influenced by the movement of the articulators that are still active from producing a preceding sound (= perseveratory or carryover coarticulation) or already active for producing the following one (= anticipatory coarticulation). With respect to syllables, it has been claimed (e.g., by Fujimura & Lovins, 1978), that there is more coarticulation within than between syllables.

Unfortunately, although there are many studies of coarticulation, there are not many studies on the effect of syllable structure on coarticulation. Researchers investigated the movements of articulatory organs, for example, the lowering of the velum for the production of nasals like [m] or [n], as a function of the nasals position in an utterance. If one found that articulatory movements differed for syllable-initial and syllable-final sounds, this could be used to define the syllable articulatorily. However, the syllable position effects reported in these studies could often also be word position effects. Browman and Goldstein simply stated that "syllable-position effects are similar to [..] word-position effects" (Browman & Goldstein, 1992, p. 166), without explicitly showing this. Turk (1994) investigated the closing and opening gestures of bilabial stops in syllable-initial (*repair*), syllable-final (*captor*), and a third position, where the stop intervocalically preceded an unstressed syllable (*leper*). In these examples, syllable-initial stops are not word-initial, and syllable-final stops are not word-final, which allows word and syllable effects to be disentangled. However, in other stimuli of her study, word and syllable position were confounded. A positive example is the study by Krakow (1989). She controlled for syllable and word context in a study of velum movements. She found that velum movements differed in word-final and word-initial positions (*hoe me* versus *home E*), where velar lowering was significantly earlier for final than for initial nasals. In pairs where the nasals occupied the same syllabic but different word positions (e.g., *pig me* versus *pigmy*), nasals in word-final and word-medial syllable coda positions showed similar velum movements. Krakow concluded that the different velum movements in pairs like *hoe me* and *home E* are caused by the different syllable positions that the nasals occupy (see also Boyce, Krakow, Bell-Berti, & Gelfer, 1990; Boyce, Krakow, & Bell-Berti, 1991). More evidence like this is needed to examine the relation of articulation and syllable structure.

Thus, to obtain coarticulatory evidence for or against syllable structure at all, the context has to be carefully controlled to disentangle syllabic from word level effects. At the moment, too little is known about the relation of coarticulation and syllabic structure to allow the former to be used to define the latter.

Instead of the articulatory properties, the acoustic properties of speech could provide a definition of syllables. In an acoustic definition of the syllable, the concept of sonority might be useful. A sound's intrinsic sonority (i.e., acoustic energy) can be phonetically measured in comparison to other sounds that have been produced under the same loudness, stress, and length conditions. As Sievers (1881) noted more than 100 years ago, the vowel in the syllable's nucleus has the highest sonority, which drops to both ends of the syllable. The segments' sonority hence raises from the start of a syllable to its nucleus - normally the vowel - and falls towards the end. This regularity has come to be known as the Sonority Sequencing Principle. Sounds can be ranked in a scale according to their sonority. Different languages can structure the sonority scale in different ways, but plosives like [p] or [t] are always on the least sonorous end and vowels are the most sonorous elements. A syllable in which a low-ranked stop follows a higher-ranked consonant, as in /lta/, is ruled out by the Sonority Sequencing Principle.

However, a sound's sonority is not a fixed measure and can vary depending on the speaker and factors like loudness, length, and stress, especially in continuous speech. Without an objective measure, it becomes hard to define sonority phonetically without referring to the positions that sounds can hold in a syllable. But it was the syllable that the concept of sonority should help to define, and this makes the argument a circular one. A further problem are exceptions to the Sonority Sequencing Principle. Many languages allow onset clusters where a fricative precedes a stop, although the stop is ranked lower on the sonority scale and should therefore be closer to the syllable's edge than the fricative. Examples are English *school* [sku:l], Dutch *spier* [spi:r] ("muscle"), French *stage* [sta:ž] ("internship"), or German *Sprung* [špr.ʊŋ] "jump". The same problem also occurs in syllable codas, where the low-ranked stop may be followed by a higher-ranked fricative. One would have to redefine the sonority scale and assign the same degree of sonority to fricatives and stops to explain the exceptions. A further problem are nasals and liquids, which sometimes can occur with syllabic status, like English *paddling* that can be produced with either two or three phonetic syllables ([pæd.lɪŋ] or [pæd.l.ɪŋ], example from Ladefoged, 1993).

Instead of referring to sonority, one could check the acoustic signal for characteristics that syllable structure imposes on the acoustic characteristics of segments. For instance, Maddieson (1984) lists evidence for Closed Syllable Vowel Shortening from many studies in different languages: Vowels are shortened before tautosyllabic consonants when compared to heterosyllabic ones. This could serve as a potential phonetic cue for syllabic structure. Rietveld and Frauenfelder (1987) found that Dutch vowels are longer when a syllable boundary intervenes between the vowel and the following consonant (CV.C), compared to the vowel preceding a coda consonant (CVC.). They furthermore suggested for Dutch that vowel duration varies not only as a function of syllable boundary, but also as a function of the following consonant, and these two factors interact: Vowels were shortened more before /m/ or /r/ than before /s/ or /l/. Nittrouer et al. (1988) investigated the coordination of articulatory movements in contrasting utterances like /pa.map/, /pam.ap/, /pa.pap/, /pap.ap/ and found shorter durations for vocalic gestures (jaw cycles) when the first syllable was closed than for open syllables. The duration of consonants in clusters is smaller than the duration of single consonants (e.g., Crystal & House, 1988). Most durational data were obtained in reading tasks in a laboratory situation and it is not clear whether these measures are useful for investigating continuous speech (for differences between read and spontaneous speech see Blaauw, 1995). Moreover, not only do many factors affect segmental durations. In addition, a large number of different segmental durations could in combination provide cues to syllabification (Anderson & Port, 1994). Furthermore, even for read speech, the results are not clear. Crystal and House (1990) did not find vowel shortening for stressed vowels in closed syllables of monosyllabic words.

Another often cited example for a phonetic cue to syllabic structure is the English lateral /l/, which is 'light' in syllable initial and 'dark' in syllable-final position. Another cue is velum height in nasals, which is said to be higher in syllable-initial than in syllable-final position in English and Japanese (Fujimura & Lovins, 1978). But while these phenomena may serve as language-specific cues for syllable structure, they cannot be used to define the syllable phonetically.

In sum, a uniform phonetic correlate for the syllable is not available, and the cues for syllabification differ across languages. Many phoneticians view syllables as epiphenomena of other phonetic events and because of its vagueness the role of the syllable in phonetic theories is limited. The evidence reported in the previous section will become relevant for Experiment 4 in Part 1 of this thesis.

The Syllable in Phonology

There is a clear preference for the syllable in phonology, although Chomsky and Halle's (1968) influential *Sound pattern of English (SPE)* neglected it as a relevant phonological unit. This early generative framework only concatenated linear strings of phonological segments and boundary symbols for syllables, morphemes and words. While rules in *SPE* could refer to syllable boundaries, they could not refer to syllabic units. However, earlier theoreticians like Trubetzkoy, Hjelmslev, Firth, or Fudge had included the syllable as a unit in their theories (for an overview see Anderson, 1985). And although *SPE* dominated the field, some phonologists like Hooper (1972) or Schane (1972) stated the importance of the syllable as an essential phonological unit.

In the meantime, the syllable has been accepted in almost every phonological theory, because it has proven to be useful in several aspects, to which I will turn below. While there is agreement about the important role of the syllable in phonology, the theories for the structure of syllables diverge. While in some theories syllables have no internal structure (e.g., Kahn, 1976; Clements & Keyser, 1983), most phonologists assume that a syllable contains subparts. Two models are particularly common, and I will describe them briefly, see Figure 1.

The first model is the traditional theory of the syllable. The theory assumes that the syllable consists of three subparts (e.g., Goldsmith, 1990; Selkirk, 1982) . The first part, the 'onset', contains the consonant(s) preceding the vowel. The second part, the vowel, is referred to as the syllable's 'nucleus', and the 'coda' finally comprises the consonant(s) following the vowel. Every syllable has to consist of at least a nucleus, as the first syllable $(\text{I:})_\sigma$ in *even*. Syllables that contain one or more coda consonants, like the syllable $(\text{ɒd})_\sigma$ in *odd*, are called 'closed syllables'. Many researchers furthermore assume that the nucleus and the coda together form a unit called the 'rhyme'. Arguments in favor for the rhyme come, for instance, from phonotactics. Spanish syllables have maximally three elements in their rhyme, regardless of the number of consonants in the onset (Harris, 1983, p. 9f). The same holds for Turkish (Clements & Keyser, 1983).

The second model (e.g., Hayes, 1989; van der Hulst, 1984; Hyman, 1985) differs in the nature of units in a syllable. Syllables consists of at least one and at most two moras, which are prosodic units. A short vowel (e.g., the vowel in *odd*) is dominated by one mora, and a long vowel (e.g., the first vowel in *even*) by two moras. A consonant that follows the vowel in a syllable may also be dominated by

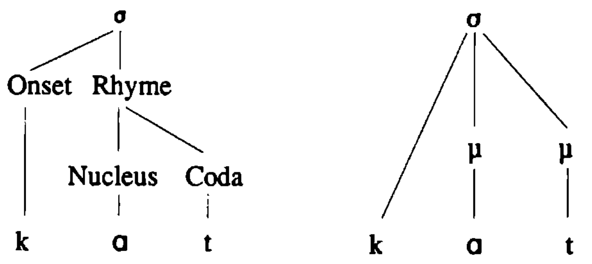


Figure 1. The Traditional Theory of the Syllable (Left) and the Moraic Theory of the Syllable (Right)
Note. The moraic syllable is depicted following the theory of Hayes (1989).

a mora, dependent on the language. In Dutch, where closed syllables with a short vowel are heavy, the coda consonant is dominated by a mora, as shown in Figure 1. Syllables with one mora are light, and syllables with two moras are heavy. The distinction of heavy and light syllables is important for the stress pattern of a word. Latin stress, for example, falls on the penultimate syllable if this syllable is either closed by a consonant or contains a long vowel, or in short, if the penult is heavy. In case the penult is light, that is, it only contains a short vowel, the antepenultimate syllable will be stressed. I will return to the Latin Stress Rule below. In contrast to nucleus and coda, the onset of a syllable does not count in determining syllable weight. Therefore, it has no mora on its own. Following Hayes, onset consonants are immediately attached to the syllable node and have therefore a different status from the other segments in the syllable, but there are alternative accounts (e.g., Hyman, 1985).

Whether the moraic or the traditional model is to be preferred is still an open question. The distinction between heavy and light syllables, for instance, which is provided by the number of moras in the moraic theory, can be accounted for in the traditional model as well: Heavy syllables have a branching rhyme, that is, either a nucleus and a coda in closed syllables, or two vowel units in the nucleus in the case of long vowels. Light syllables contain only a nucleus with a short vowel. For an evaluation of the two alternative accounts see, for instance, Blevins, 1995; Kenstowicz, 1994; or Pierrehumbert and Nair, 1995.

Although the models differ, they all agree that syllables are needed to account for phonotactic constraints, domains of phonological rules, and phenomena like epenthesis and deletion (see the arguments in Clements and Keyser, 1983; Kahn, 1976; Selkirk, 1982, to name just a few).

Referring to the syllable, phonologists can explain phonotactic constraints. For instance, certain segments or clusters are not allowed in syllable onsets, like /ŋ/ or /dn/ in Dutch. The non-existence of some clusters can be explained by referring to the Sonority Hierarchy (see, e.g., Clements, 1990).

The syllable furthermore helps formulating domains for phonological rules. These rules can either refer to the syllable as a whole, or to certain positions within the syllable, or to the structure of a segment's syllable. First, the classic Latin Stress Rule mentioned above refers to the syllable as a whole unit. The rule states that penultimate syllables are stressed if they are heavy. Otherwise stress falls on the antepenultimate. In *SPE*, which could not refer to the syllable at all, this rule was very complicated. Rules in SPE could only refer to the boundaries of morphemes and words. But even if a syllable boundary had been included in SPE as a further type of boundary, the resulting rule would have been still complicated enough:

$$V \rightarrow [+\text{stress}] \,/\, \_((C_0[V, -\text{long}]\$)C_0V)C_0\#.$$

In words, the rule says that a vowel is stressed if it stands before zero or more consonants followed by a short vowel followed a syllable boundary followed by zero or more consonants followed by any vowel followed by zero or more consonants. This is the case, for instance in words like *dómina* "mistress", *cellárius* "cellarer", or *distántia* "distance". If these conditions do not hold, a vowel is stressed if it stands before zero or more consonants followed by any vowel followed by zero or more consonants (like in *dómus* "house", or *ménsa* "table"). The Latin Stress Rule can be formulated easily as soon as one refers to the syllable as a unit: "Stress the penult if it is heavy, otherwise stress the antepenult". No doubt - the syllable is a useful unit to formulate phonological rules.

Second, not only the syllable as a whole unit, but also certain positions within the syllable can condition phonological rules. Syllable-final stop consonants and fricatives devoice, for instance in Dutch or German, and syllable-initial obstruents are aspirated in English.

A third type of phonological rule refers to the structure of a segment's syllable. Québécian French, for instance, has a rule that makes high vowels, like /I/, lax in closed syllables. While the masculine form of *petit* ("small") is [pt^{s}i], the female form *petite* has a closed syllable and hence a lax vowel: [pt^{s}ɪt].

The syllable thus helps to define the domain for all three types of phonological rules. In addition, processes at the segmental level can be explained as attempts to obey certain syllable template conditions. Take, for instance, English roots like *hymn* or *damn*. /mn/ is not a legal English coda cluster, because the sonority of segments within a syllable coda should decrease according to the Sonority Sequencing Principle, as mentioned in the section on the Syllable in Phonetics. /m/ and /n/, however, are of similar sonority and hence do not obey the Sonority Sequencing Principle. Consequently, the final /n/ cannot be included in the syllable template and is therefore not produced. As a result, we get a form like /dæm/. In *damnation*, where a V-initial suffix follows the stem, the /n/ associates to the onset of the second syllable and therefore surfaces. Now consider the word *meter*. Its stem contains the illegal coda cluster /tr/. As in *damnation*, the final /r/ associates to the following onset of a V-initial suffix, as in *metrical*. When no suffix follows, however, the word is not produced as /mi:t/, but the /r/ vocalizes and surfaces as a second syllable.

In addition to different models for syllabic structure, different mechanisms have been suggested for syllabification in phonology. Syllables are constructed either by rule, or by template mapping. Rule-based approaches differ in detail. But all of them take a string of segments and first assign a syllable nucleus to the vocalic elements. For instance, syllabifying the segments of the word *antique* /ænti:k/, /æ/ and /i:/ become syllable nuclei, and each of them forms the core of a syllable. The word has hence two syllables to which the remaining consonants are distributed. Consonants are assigned to the following syllable if the result is a legal syllable onset (= Maximal Onset Principle). It is a common phenomenon in the world's languages that consonants prefer to be in the onset of a syllable. This will not work for the final /k/ in *antique*, since no syllable follows it. But the /t/ will associate to the onset of the second syllable. The /n/ cannot do that, because /nt/ would offend the Sonority Sequencing Principle. Consonants that cannot become onsets go in the coda of the preceding syllable, and the syllabified form of *antique* is $(\text{æn})_\sigma(\text{ti:k})_\sigma$.

Alternatively to syllabification by rule, syllables are built by template mapping: a string of segments is matched to a template that functions as a well-formedness condition. Some languages, for instance, Hawaiian or Samoan, do not allow coda consonants. Wellformed syllables have to be open. However, sometimes foreign loanwords enter from a language that allows different syllable templates. Samoan syllables, for instance, contain a vowel that may be preceded by one optional consonant. English words have a more complex syllable template. To prohibit closed syllables, a vowel may be added to a word that ends in a consonant, like in Samoan *vulu* "wool", or *uniti* "unit". Due to the final vowel, the English coda consonant (e.g., the /l/ in *wool*) becomes a consonant preceding a vowel and thus obeys the syllable template. Consonant clusters can also be split up by intervening vowels like the second /i/ in *piniki* "pink". Instead of adding a vowel, consonants may be deleted to suffice the Samoan syllable template, for instance, the /k/ in *iniseti* "insect", or /p/ and /b/ in *setema* "September" (the Samoan words are taken from Cain, 1986).

In many languages, the Maximal Onset Principle neglects intervening morphological boundaries. As a result, syllabic and morphological boundaries do not align, like in *grounding*, which syllabifies /graʊn.dɪŋ/, though its morphemes are *ground+ing*. In some languages, even word boundaries are neglected, as Harris' often cited example from Spanish shows (Harris, 1983, p. 43):

(1)

Syllabification of the single words:
Los o.tros es.ta.ban en el a.víon

Resyllabified form (bold segments=new onsets)
Lo.**s**o.tro.**s**es.ta.ba.**n**e.**n**e.**l**a.víon

"The others were on the airplane"

I will use the traditional term 'resyllabification' for this phenomenon. A resyllabified consonant delinks from the coda of the first syllable and associates to the onset of the second syllable. Although some phonologists argued to replace resyllabification by other mechanisms (e.g., Itô, 1986; Rice, 1990), I will stay with the traditional notion of resyllabification (see the General Discussion for more details).

What are the implications of this section for the experiments presented below? Levelt's model of phonological encoding is compatible with both the moraic or the traditional theory of the syllable. I will use the traditional terminology throughout this thesis and refer to syllable onsets, nuclei, and codas. Syllables are constructed in the model not by template mapping, but by segment-to-frame-association rules. The rules make a vowel a syllable nucleus and a consonant an onset, if this results in a legal onset. The issue of resyllabification, especially between words, will be important in Part 2 of the thesis, because the model of phonological encoding does not assume resyllabification. Instead, a speaker immediately produces the syllables that surface in connected speech.

The Syllable and Syllabification in Dutch: Single Words

This is the moment to look at the syllable inventory of Dutch, the language tested in all experiments (see Booij, 1995; van der Hulst, 1984). Some general features hold for all Dutch syllables. A syllable needs a vowel as a nucleus (although some dialects allow for syllabic nasals), which can be either short or long or a diphthong. Zero to two segments fill the onset. A rhyme contains two or three positions, which can be filled by a long vowel (= two positions, e.g., *vla* /vla:/ "cream") or a long vowel and a consonant (= three positions, e.g., *kaas* /ka:s/, "cheese"), or a short vowel and one consonant (= two positions, e.g., *bus* /bʏs/ "bus") or two consonants (= three positions, e.g., *drank* /drɑŋk/ "drink"). In addition, coronal consonants can be added at the end of words to build extra long consonantal clusters, like for instance in the words *herfst* ("autumn") or *promptst* ("most prompt"). The coronal consonants are interpreted as appendices to the prosodic word, which will be defined below. A word can start with three consonants, if the edge-most consonant is the coronal fricative /s/, like in *split* ("split"). There is some discussion in the literature on the status of the initial coronal fricative. It is either seen as an initial appendix to the prosodic word (Trommelen, 1984), or as a part of the syllable onset (van der Hulst, 1984, Booij, 1995).

Within the domain of syllabification, Dutch as many other languages prefers to have as many consonants in syllable onset position as possible, as long as they obey the Sonority Sequencing Principle. Each monomorphemic word, and each constituent of a compound is a domain of syllabification: *[koop]$_N$[avond]$_N$* ("shopping night") does not syllabify *koo.pa.vond*, as the Maximal Onset Principle would predict, but *koop.a.vond.* Moreover, prefixes are domains of syllabification: $[ont]_{PREF}[erven]_V$ ("disinherit") syllabifies as *ont.er.ven*, not *on.ter.ven*. Finally, some suffixes, are domains for syllabification. They are all consonant-initial (exception: the vowel-initial *-achtig* "-ish") and have at least one full vowel. For instance, $[verf]_N[loos]_{SUF}$ ("paintless") does not syllabify *ver.floos*, but *verf.loos*. If a suffix is vowel-initial, thus does not build a separate domain of syllabification, a preceding consonant will occupy its onset position. For instance, $[boot]_N[en]_{PLURAL}$ ("boats") syllabifies *bo.ten.* If the syllable of which the consonant was the coda is left with only a short vowel in the nucleus, the consonant will be ambisyllabic, since short open syllables are not allowed in Dutch. The ambisyllabic consonant will simultaneously belong to the coda of the first and the onset of the second syllable, like the /k/ in the plural of *rok* /rɔk/ ("skirt"), which is *rokken* /rɔ(k)ən/ to avoid an illegal short open first syllable.

The domain of syllabification in Dutch coincides with the prosodic word (ω) in most cases. This prosodic constituent has as one defining feature its behavior in coordination reduction (Booij, 1985): In these reductions, parts of complex words are deleted when there is another complex word with the identical part in the same phrase, like in *minimum(bedragen) en maximumbedragen* "minimum (amounts) and maximum amounts". The deleted part cannot be defined on the basis of syntactic or morphological information. It can be a part of a compound, while the surviving part is not, and vice versa, as in *ijs(beren) en bruine beren* "polar bears and brown bears". The deleted parts can be nouns, adjectives, prepositions, verbs, or quantifiers. Some suffixes delete, too: *zwanger(schap) en moederschap* "pregnancy and motherhood". But not all suffixes delete: *absurd(iteit) en banaliteit* "absurd(ity) and banality" is not possible. The unifying feature is that all deleted parts are prosodic words.

Furthermore, Dutch has a rule of schwa deletion that has the prosodic word as its domain. The final schwa in *zijde* [zɛidə] ("silk") deletes when the suffix *-ig* [əx] is attached to it: *zijdig* [zɛidəx] ("silky"). In compounds, however, and before some affixes, the schwa is not deleted: *zijdeinkomsten* [zɛidəɪŋkɔmstən] ("silk revenues"), *zijdeachtig* [zɛidəɑxtəx] ("silk-like") (Booij, 1995).

Prefixes have a special status. Since they do not delete in coordination reduction and mostly contain no full vowel, they are not prosodic words on their own. On the other hand, they do not undergo the rule of schwa deletion, like in *geacht* [ɣə.ɑxt] ("respected"), which does not become [ɣɑxt]. Prefixes hence differ from suffixes, but behave like a member of a compound. Booij (1985, 1995) therefore argued that prefixes are chomsky-adjoined to the prosodic word. This results in a structure like $[(\text{bə})_\sigma\ [(\text{a:})_\sigma(\text{dəm})_\sigma]_\omega]_\omega$.

Syllable and prosodic word are units of a larger prosodic hierarchy, which differs between theoretical variants, but may include from bottom to top the syllable, foot, prosodic word, clitic group, phonological phrase, intonational phrase, and the utterance (this is the hierarchy proposed by Nespor and Vogel, 1986). Of all prosodic constituents, only syllable and prosodic word will be of relevance in this study. Foot structure will not play a role, since it was not varied in the experiments.

Syllabification Across Word Boundaries: Encliticized Forms in Dutch

In Dutch, many function words (pronouns, determiners, conjunctions, etc.) have two forms, a phonologically strong and a phonologically weak one. The phonologically strong form contains a full vowel (e.g., *het* "it" /hɛt/), while the corresponding weak form (henceforth 'clitic') normally has only schwa as a vowel (e.g., *het* "it" /ət/). Examples for strong and corresponding weak forms are given in (2):

(2)	strong	weak		
	wɛɪ	wə	*wij*	"we"
	ɛn	ən	*en*	"and"
	ens	əs	*eens*	"once"
	hɛt	ət	*het*	"it / the (neuter)"
	hθn	zə	*hun*	"them (dative)"
	jɑu	jə	*jouw*	"your (sg.)"
	mɛɪ	mə	*mij*	"me"

The weak forms have to be listed as variants in the lexicon, because there are no general phonological processes in Dutch by means of which the weak forms could be derived from the strong forms. The weak determiner *de* (/də/ "the") does not even have a corresponding strong form. Furthermore, some weak forms differ from the corresponding strong forms in meaning. For instance, the weak form *ze* (/zə/ "they") can refer to persons and things, while the strong form *zij* /zɛɪ/ can only denote persons. Also certain idioms only allow the weak form, for instance, *laat me / *mij niet lachen* ("don't make me laugh") (Berendsen, 1986). Whether a speaker decides to use the weak or the strong form in cases where both are allowed syntactically and semantically, depends among other things on the speech tempo. In careful, slow speech, the full forms are more likely, while in fast speech, the weak forms will be used.

Note that prosodic words never start with a schwa in Dutch and never have exclusively schwa as a vowel (Booij, 1996), while many of the weak forms show these properties. To avoid a schwa-initial syllable, and in accordance with the general tendency of languages to avoid syllables that lack an onset, a schwa-initial weak form of a function word will prefer having the coda element of the preceding word in its onset position, resulting in an encliticized form, as shown in (3):

(3)	$(\text{drɪŋ})_\sigma(\text{kət})_\sigma$	*drink het*	"drink it"
	$(\text{daŋ})_\sigma(\text{kəm})_\sigma$	*dank hem*	"thank him"
	$(\text{bo:})_\sigma(\text{tən})_\sigma$	*boot en*	"boat and"

Since the clitic has exclusively schwa as a vowel, it cannot be accented or contrastively stressed, and it never topicalizes. The fact that the clitic does not fulfill the requirements on a prosodic word speaks against an analysis of *drink het* as a clitic group (for the clitic group, see Nespor & Vogel, 1986). A clitic group has to consist of prosodic words, since it is widely assumed that the prosodic hierarchy is strictly layered: Constituents of one level are exhaustively parsed into constituents of the level immediately dominating it. Dutch clitics, however, are not prosodic words on their own. I will follow the analyses by Booij (1996), Gussenhoven (1985), Lahiri et al. (1990), and Neijt (1984), and assume that clitics are prosodically integrated into an adjacent prosodic word. The clitics investigated in this thesis are all enclitic, that is, they integrate into the preceding prosodic word. This is the preferred direction of cliticization in Dutch (Booij, 1995, Gussenhoven, 1985).

Booij (1996) argues for the integration of enclitics into the preceding foot and treats them as similar to suffixes. For this, he has to allow for ternary feet. Languages seem to prefer feet that contain two syllables (= 'binary feet'). When a clitic incorporates into a preceding foot that already has two syllables, the result will be a ternary foot, like in *kochten het* "bought it": $[((\text{kɔx})_\sigma(\text{tə})_\sigma(\text{nət})_\sigma)_{\text{FOOT}}]_\omega$. Whether the enclitic is incorporated into the preceding foot, or alternatively chomsky-adjoined to the preceding foot like in $[(((\text{kɔx})_\sigma(\text{tə})_\sigma)_{\text{FOOT}}(\text{nət})_\sigma)_{\text{FOOT}}]_\omega$ - in which case we would have recursive feet, or whether it is immediately linked to the prosodic word, skipping the foot level as in $[((\text{kɔx})_\sigma(\text{tə})_\sigma)_{\text{FOOT}}(\text{nət})_\sigma]_\omega$, is not crucial for my argument. Important is that the clitic belongs to the preceding prosodic word. This claim can be defended, since the rules that apply obligatorily within prosodic words also apply obligatorily in host+clitic combinations. An example is the rule of homorganic glide insertion to avoid hiatus in Dutch, as in the word *duo* [dʏ:ɥo], *Ruanda* [ruʋɑnda], or *bioscoop* ("cinema") [bijɔsko:p] (Gussenhoven, 1980). This rule also applies to host+clitic combinations: *zie het* ("see it") [si:jət], *doe het* ("do it") [duʋət]. More evidence comes from the optional rule that deletes /n/ syllable-finally after schwa (e.g., *lezen* "read" /le:zə/ or /le:zən/). This rule is blocked in encliticized forms, where the /n/ has to appear: *zij lezen het boek* ("they read the book") *zij* /le:zənət/ *boek*. The /n/ is no longer in coda position, but in the onset of the schwa-initial clitic syllable and may therefore not delete (Booij, 1995).

Encliticized forms will be relevant in both parts of the thesis, because they play a major role in Levelt's model of phonological encoding, as I will describe below. Encliticized forms serve as targets in the syllable priming experiments that are reported in Part 1 of this thesis. Part 2 looks at the relation of encliticization and syllable-final devoicing.

The Syllable in Psycholinguistics

Finally, I will turn to psycholinguistics, where an increasing interest in the syllable has evolved in the last two decades. This section deals with studies on the syllable's role in early language acquisition and in language perception, and reports experimental investigations of syllabic structure. Studies on the production of syllables will be discussed in a separate section.

Experimental evidence for the syllable's intuitive plausibility comes, for instance, from studies with illiterates, who can segment an incoming sequence of speech in syllables more easily than identify its phonological segments (see, for instance, Morais, 1985). Phoneme identification clearly correlates with alphabetic literacy - Portuguese illiterates (Morais, Cary, Alegria, & Bertelson, 1979) or non-alphabetic literate Chinese had great difficulties to identify phonemes. Preliterate children performed substantially worse than first graders in a task where they had to indicate the number of segments in a spoken word by an equivalent number of taps (Liberman, Shankweiler, Fischer, & Carter, 1974). However, the awareness of syllables is not completely independent from literacy, either. Preliterate children performed also worse than first graders in tapping syllables. However, for syllables they performed much better than for phonemes, while the first graders performed only slightly better for syllables than for phonemes. These results suggest that syllabic awareness is less dependent on literacy than phonemic awareness.

Only recently, researchers have become interested in the role of the syllable in language acquisition, both with respect to children's early production and comprehension skills. It has often been stated that children start producing CV-sequences with adult-like timing at about six to eight months age (see, e.g., the references in Oller and Lynch, 1992; Vihmann, 1992). Certain syllables seem to be more accessible to infants than others. For instance, [hɑ] and [hə] occur frequently in early babble productions, also for French children, who do not find them in their input (Vihman, 1992). This indicates an articulatory, neuromotor source rather than a phonological role of the early syllables in babbling (see also Macken, 1992).

At a later stage, children try to maintain metrical patterns in their early productions. The distinction between strong and weak syllables is thus important. It has been argued that English as well as Dutch children preferably produce metrical patterns that consist of a strong and an optional following weak syllable (e.g., Gerken, 1991, 1994; Wijnen, Krikhaar, & den Os, 1994). For instance, English children will drop the initial weak syllable in *giraffe* more likely than the final weak syllable in *monkey*, where in the latter form the weak syllable follows the strong one. This may result from the fact that most metrical feet in English are trochees (Selkirk, 1980) and children try to align the trochaic template with strong syllables in the utterance. Gerken (1994) moreover suggested that children are in addition sensitive to higher level prosodic phrasing and map the trochaic metrical template only within higher level prosodic units in an utterance.

As has been shown by Jusczyk, Cutler, and Redanz (1993), preference for certain metrical patterns does not only show up in early productions. Nine-month-old English infants preferred the trochaic stress pattern in speech perception, too, even when the input was low-pass filtered, such that only the prosodic structure, but not the segmental content of the input was available.

Not only the distinction between strong and weak syllables, but also phonotactic constraints for syllables could be important in perception, where infants have to segment the continuous speech signal into discrete words and to discover the syntactic organization of utterances. Traditionally, many models of language acquisition start at a point where the child has already segmented the continuous signal into words and assume that the child develops a syntactic or semantic representation based on these lexical units. Stimulated by prosodic phonology and linguistic research on the syntax-phonology-interface, acquisition researchers have also considered the possibility that children use the prosodic phrasing of utterances as cues to syntactic structure. Infants showed to be sensitive to prosodic phrasing of an utterance, which was induced by inserted pauses (e.g., Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Gerken, Jusczyk, & Mandel, 1994).

While the prosodic units that help to identify syntactic clauses and phrases are larger than the syllable, syllables may help to segment the input into words. For instance, infants may use phonotactic constraints on syllables, like the aspiration of syllable-initial stops in English. For example, the absence of aspiration for the /t/ in *night rate* indicates that the stop is in coda position. According to the Maximal Onset Principle, it should build an onset cluster with the following liquid. One can conclude that an intervening word boundary blocks the onset maximization. When no word boundary intervenes, like in *nitrate*, the stop is aspirated and released. Children may use these phonotactic cues to word boundaries. Various studies have shown that they can make subtle distinctions between speech sounds (for a summary see, e.g. Jusczyk, 1992). Whether they use this ability to identify word boundaries based on syllabic wellformedness constraints, however, remains to be shown.

In sum, phonotactic regularities that are defined with reference to syllable structure and the fact that the syllable bears stress may both be relevant for early language acquisition. The phonotactic regularities may provide cues for word boundaries during speech segmentation, but whether they really do so has still to be proven. Stress is relevant for metrical structuring, for which infants showed sensitivity, both in early productions and perception. However, the relevant prosodic unit was the metrical foot, and the syllable is only important in so far as it constitutes feet. More research is needed for determining the role of the syllable as a phonological unit in language acquisition.

Turning from infants to adults, researchers have investigated how listeners may use the syllable in language perception. Mehler, Dommergues, Frauenfelder, and Segui (1981) found that the syllable plays a role in speech perception for speakers of French: In a syllable detection task, a target like *pa* was detected faster in *pa.lace*, where it corresponds to the first syllable, than in *pal.mier*, where it does not, while *pal* was detected faster in *pal.mier* than in *pa.lace*. Bradley, Sánchez-Casas, and García-Albea (1993) replicated the interaction of target syllable and carrier word's syllabic structure for Castilian Spanish. According to Mehler and other researchers, the syllable effects are evidence for stored syllables that are used as units of perception at a prelexical level to mediate between the incoming speech signal and the lexicon.

Pallier, Sebastian-Gallés, Felguera, Christophe, and Mehler (1993) found syllabic effects in phoneme detection and phoneme decision tasks for Spanish and French. The targets were phonemes that shared their segmental position, but differed in their syllabic position, for instance, *B* in *no.Blesse* and *suB.til.* Participants completed a list that included many fillers which were designed to induce an attention shift to a specific syllabic position: One group of participants received a majority of fillers where the target phoneme was in the coda of the first syllable. For a second group it was in the onset of the second syllable most of the time. All participants got the same target words and they detected target phonemes faster when they occurred in the syllabic position that had been dominant in the list. Thus, participants whose attention had been shifted to the coda of the first syllable responded to *B* faster in *suB.til* than in *no.Blesse*, and participants who had more fillers with the target phoneme in the onset of the second syllable in their list were faster for *B*s in the latter word. This showed that listeners are sensitive to syllable structure in this phoneme monitoring task.

However, the role of the syllable in speech recognition may be different crosslinguistically. Cutler, Mehler, Norris, and Segui (1986) did not find effects similar to those obtained by Mehler et al. (1981) for speakers of English, a language that has less clear syllable boundaries due to many ambisyllabic consonants. The authors assumed two steps in comprehension, namely segmentation and classification, and for the former process, everything that is salient in a language will be used to set points in the signal where the classification process can start. The salient points are syllables in the case of French, but strong syllables in the case of English. Bradley et al. (1993) did not find the interaction for English speakers that they found in Spanish. Spanish speakers did not show the interaction either when monitoring English materials.

The results obtained by Pallier et al. (1993) were partly replicated by Protopapas, Finney, and Eimas (1995) for English. In their experiments, participants monitored for the third segment of a target word, which could occur in the coda of the first syllable (e.g., the /g/ in *magnet*), or in the onset of the second syllable. For items with initial stress, this onset segment was ambisyllabic (e.g., the /g/ in *juggler*), whereas for items with second-syllable stress it was a clear coda segment (e.g., the /k/ in *seclude*). The syllable effect only occurred for targets that were stressed on the second syllable, and had clear syllable boundaries. Syllable effects thus can be obtained for English listeners in this phoneme monitoring task, if ambisyllabicity is controlled for.

For Dutch, Zwitserlood, Schriefers, Lahiri, and van Donselaar (1993) found in a syllable monitoring task a similar effect that had been found for French by Mehler et al. (1981). CVV-targets were detected faster than CVVC-targets in CVV.CVC-words, while the opposite was the case for CVVC-words. For instance, /ma:/ was detected faster in *magen* /ma:.ɣən/ ("stomachs") than in *maag* /ma:ɣ/ ("stomach"), while the opposite was the case for /ma:ɣ/. However, participants were faster in detecting strings that were not syllables, but morphemes, like /ma:ɣ/ in *magen* than strings that were neither syllables nor morphemes, like /ma:ɣ/ in *maagd* ("virgin"). This difference was less pronounced for targets that were syllables and morphemes, like /ma:ɣ/ in *maag* compared to targets that were syllables, but no morphemes, like /ma:ɣ/ in *maagden* ("virgins"). Although this difference did not reach significance, it still was a trend indicating that other confounded factors influenced the detection latencies than only syllable match.

McQueen and Fox (1995) had Dutch participants detect words that were embedded in nonwords, like *rok* ("skirt") in *fim.rok* and *fi.drok*. Participants responded faster when the target corresponded to a syllable than when it was part of a syllable. The syllable boundaries were determined by phonotactics. In the example above, /mr/ in *fimrok* is not a legal cluster, and the voiced /d/ in *fidrok* can only be an onset, because it would devoice in coda position. Further research has to show how participants use phonological constraints on syllable structure to segment continuous speech.

There is thus evidence that the syllable plays a role in speech perception. The importance of syllabic structure may be greater when the language favors syllables of a particular type, like open syllables in French or Spanish.

Treiman and colleagues investigated the role of the syllable in metalinguistic tasks. In word game tasks they found evidence for the split of the syllable in an onset and a rhyme that many phonologists have proposed for independent reasons. When participants had to blend two nonsense monosyllables with onset and coda clusters into one syllable, they combined the onset cluster of one syllable with the rhyme of the other syllable in 67% of the responses, while the rest were blends splitting the two consonants of the onset cluster (10%), the consonants of the coda cluster (7%) or the nucleus and the coda (6%) (Treiman, 1983). When they had to blend the two syllables according to blending rules they had learned, the rule of combining an onset cluster with a rhyme caused the smallest amount of errors compared to the rule that divided the onset cluster, or as compared to the rules that divided the syllables within the rhyme constituent. Fowler, Treiman, and Gross (1993) used a phoneme shift task: Participants saw two stimuli and had to shift one or more phonemes that were capitalized from one form to the other. For instance, they had to respond with the form *kepniz* to the stimuli *bepniz Kugfam*. It turned out that shifting the whole onset was faster than shifting more or fewer phonemes than the onset. One might ask whether the visual presentation of nonword stimuli permits the investigation of the role of syllables in speech production. But the authors also used a word game task where participants were instructed to replace an onset or coda consonant of the medial stressed syllable of an auditorily presented three-syllable nonword by always the same consonant. Manipulating the onset was again easier than exchanging the coda.

Treiman (1984) found that American English liquids show more cohesion to the vowel in the nucleus than obstruents: While e.g. blending of *als* and *ord* led to VC/C blends in half of the cases, for *ooks* and *eft* the majority of responses were of the V/CC-kind. This is an interesting finding, since it provides some empirical evidence for the importance of the sonority scale, at least in syllable codas. A potential problem is that the American English sequence *ar* in *ard*, which results from blending *als* and *ord* is phonetically speaking only one sound, namely a vocalized /r/, and this may favor blending. However, this does not hold for the /l/.

Treiman and Danis (1988a) were interested in ambisyllabicity and looked at polysyllabic words, using a syllable reversal task. Participants got words like *melon*, *relate*, *topic* and *guitar*. Their task was to say the syllables in backward order. For *melon*, possible answers were *on mel*, indicating that /l/ was treated as the coda of the original first syllable, *lon me*, where the liquid was the onset of the original second syllable, and *lon mel*, indicating that it was ambisyllabic. The results indicated that obstruents had the tendency to be in onset position of the original second syllable, like *pic to*, while liquids were in the coda of the original first syllable, like *on mel*. However, when the second syllable was stressed, the liquid was likely in the onset of the original second syllable, like *late re*. This possibility was still more likely for obstruents, like *tar gui*. Finally, consonants were more likely in the coda of the original first syllable after a short than after a long stressed vowel. Thus stress, length, the type of consonant, and its sonority seem to be important variables for syllabic grouping, and this is compatible with the predictions of phonological theory and phonetics.

The role of the syllable in phonological memory was also investigated. Participants who got a list consisting of six nonsense CVC-syllables at a rate of one syllable per second and had to recall them in the right order immediately after hearing them, produced most frequently the error type in which onsets were taken from one syllable and rhymes were taken from another syllable, whereas CV/C-blends occurred less frequently. Syllables seem to be stored in short term memory including intrasyllabic units (Treiman & Danis, 1988b). Bruck, Treiman, and Caravolas (1995) had participants compare two nonwords that were auditorily presented with a short intervening pause and had them decide whether those contained the same sounds. When the shared sounds constituted an entire syllable, participants showed faster detection times than when the same number of shared sounds did not constitute a syllable. This again suggests that syllables are relevant units for phonological memorization and comparison, at least for nonwords.

In sum, the work above constitutes psycholinguistic evidence for linguistic theories on syllabic structure, favoring accounts that assume subsyllabic units like onset and rhyme, or accounts that at least assign a special status to the onset consonant, contrary to theories with completely flat syllable structures.[1] Furthermore, they suggest a higher degree of cohesion between nuclei and more sonorous coda consonants than less sonorous ones. Moreover, the syllable probably serves as a unit in phonological memory tasks, and at least in some languages as a unit for on-line speech perception. The studies are important for this thesis, since they show that listeners use the syllable in language comprehension. Accordingly, it should play a role in the production of speech.

[1]This interpretation has been challenged recently by Pierrehumbert and Nair (1995), who argued that the results occur because participants align stressed vowels of two words they have to blend.

The Production of Syllables

Phonological Encoding in Speech Production

The production of syllables is one aspect of the encoding of a word's phonological form, which is in its turn part of the general issue of how words are accessed from the lexicon when we speak. In all models of language production that have been developed in the last two decades (for instance, Fromkin, 1973; Garrett, 1975; Shattuck-Hufnagel, 1979; Butterworth, 1983; Dell, 1986, 1988; Levelt, 1989), lexical access consists of two parts. This distinction corresponds to the two levels that are involved in formulating an utterance: The generation of grammatical and of phonological structure. The input for grammatical encoding is the intended prelexical message. During grammatical encoding, appropriate lexical items are selected that are not yet phonologically specified, and semantic and syntactic relations are created. Models differ as to how syntactic structure is generated. In some models, the generation is based on syntactic frames that are produced on the basis of a preverbal conceptual message and then filled with the appropriate lexical items (for instance, Garrett, 1975, 1988; Shattuck-Hufnagel, 1979). In other models syntactic encoding is lexically driven (Kempen & Hoenkamp, 1987; Levelt, 1989): The retrieval of a lexical unit, called a 'lemma', triggers the creation of syntactic structure. In both accounts, the output of the first part of lexical access is a syntactic surface structure that has to be phonologically specified. This is the job of the second part of the access process, phonological encoding.

Evidence for the distinction of grammatical and phonological encoding comes from different sources. For instance, word exchanges in speech errors involve words of the same grammatical category that often belong to different syntactic phrases, while sound errors involve segments of adjacent words irrespective of word class. The former errors probably arise at the level of grammatical encoding, while the latter arise in phonological encoding. The tip-of-the-tongue (TOT) phenomenon (e.g., Browman, 1978; Brown, 1991; Brown & McNeill 1966) provides further evidence for the distinction. Speakers in a TOT-state cannot produce the word they are looking for, although they normally know its meaning. The temporal order of the two components in lexical access is still under debate. Serial models assume that grammatical encoding precedes phonological encoding. In the strictest formulation, only the lemma that is actually selected in grammatical encoding is phonologically encoded (Levelt, 1989, 1992; Levelt et al., 1991). In other accounts (e.g., Dell, 1986; Dell & O'Seaghdha, 1992), the phonological forms of several active lemma candidates are activated before the appropriate lemma is selected and the activated phonological forms can in turn influence the encoding process on the higher levels.

For summaries on lexical access and its first stage, grammatical encoding, the reader is referred to the summaries in Bock and Levelt (1994); Fromkin (1993); Garrett (1988), or Levelt (1992). The second stage, phonological encoding, will be discussed in more detail now.

Models of speech production have been heavily influenced by speech error data (Shattuck-Hufnagel, 1979; Dell, 1986) and more recently by experimental results (Dell, 1988; Levelt, 1989, 1992). Speech error data are considered to be relevant, because it is assumed that the way a system breaks down can provide insight into the way it works. The data are either naturally occurring errors, or errors that were elicited experimentally (e.g., Baars, Motley, & MacKay, 1975). The key findings were the following. Most phonological speech errors involve segments rather than features or syllables. Segments are exchanged, deleted, added, substituted or shifted. The direction of the latter types of errors can be anticipatory or perseveratory. Examples are given in Table 1, they are taken from Fromkin (1973).

error type	error	target
sound exchange:	heft lemisphere	left hemisphere
	god to seen	gone to seed
anticipatory shift:	a meeting arathon	an eating marathon
perseveratory substitution:	gave the goy	gave the boy

Table 1. Examples of Speech Errors

The fact that the results of such errors are phonologically well-formed sequences of segments has been taken as evidence for their location at a phonological rather than a phonetic level of processing. Recent phonetic electromyographical investigations of experimentally induced speech errors, however, revealed remnants of motor activity for sounds that had been deleted in speech errors. Similarly, sounds that had been added in addition errors differed in their motor activity from similar sounds that did not result from errors (Mowrey & MacKay, 1990). The authors assume that sound errors arise on an articulatory rather than a phonological level, but an alternative hypothesis is that the errors occur at the phonological level and affect articulatory planning. Whether the sound errors arise on a phonological or articulatory level is still an open issue (for a detailed evaluation of the speech error literature see Meyer, 1992).

The occurrence of sound errors has lead researchers to believe that the phonological forms of words are not stored in the mental lexicon as syllabified units. If syllabified forms were stored, we would expect more errors that involve whole syllables as units, e.g. syllable exchanges. Those, however, are rare. Speakers instead seem to retrieve independently a word's segments and slots specifying word or syllable positions for syllables or words, and then combine the segments with the slots. When this process goes awry, we encounter a speech error. Speech errors normally involve only one segment. But when on occasion a consonant and a vowel slip together, they are usually a rhyme and not an onset-nucleus cluster (Nooteboom, 1969; Shattuck-Hufnagel, 1983), which again can be seen as evidence for the onset-rhyme split.

Many of the models have been influenced by what has been called the 'syllable-position constraint' in speech errors: Segments involved in exchange errors, shifts, or substitutions are likely to share their syllabic position. For instance, onsets exchange with onsets, or codas with codas. However, the syllable position constraint has come under some debate since it has been noticed (Shattuck-Hufnagel, 1987) that most segmental errors arise in word onset positions. The reported syllable position effect could thus also be a word onset effect.

One of the drawbacks of speech error data is that it does not provide evidence for the time course of phonological encoding. In the meantime, experiments using several paradigms have shown that phonological encoding of a word proceeds from left to right. Meyer (1991) used an implicit priming paradigm in which participants had to respond to a visually presented word by naming targets that were semantically associated to that word. The targets were grouped in blocks. In heterogeneous blocks, there was no phonological relation between the targets. In homogeneous blocks, the targets had a certain phonological overlap. Reaction times were faster in the homogeneous blocks, but only under certain conditions: For mono- and bisyllabic words, information about the onset of a target speeded up reaction times, while information about the rhyme (without shared onset information) had no facilitatory effect. Information about the syllable's onset+rhyme had a stronger effect than information about onset+nucleus. Using the same paradigm, Meyer (1990) showed that information about the second syllable of a trisyllabic word only facilitated the response when information about the first syllable had also been given. Information about the target word's first syllable always speeded up the participant's reaction times - independent of whether the second syllable was given. For bisyllabic items, the first syllable speeded up reaction times as well, while the second syllable alone did not shorten the response latencies. This suggests that segments within syllables as well as consecutive syllables are constructed from left to right.

Meyer and Schriefers (1991) found evidence supporting Meyer's claims, using a picture-word interference task. In these tasks, reaction times for naming a picture are the dependent variable used to investigate the effect of several kinds of interfering stimuli (IS) that are presented at different points in time with respect to picture presentation. The IS may exhibit various kinds of relationships to the target. In the study by Meyer and Schriefers, auditorily presented IS were either phonologically related to the beginning of words or to the end. In the beginning-related conditions, the IS overlapped in onset and nucleus for monosyllabic targets (like IS *veen* "peat" for target *veer* "feather") or in the first syllable for bisyllabic targets (IS *havik* "hawk" for target *hamer* "hammer"). End-related IS corresponded to the target word in nucleus and coda (IS *peer* "pear" for target *veer*) or the second syllable (IS *zomer* "summer" for target *hamer*). Beginning-related IS yielded facilitatory effects (compared to a silence condition) when the IS were presented 150 ms before, simultaneously with, or 150 ms after the picture. End-related IS facilitated responses only when presented simultaneously with the picture or 150 ms after the picture.

Wheeldon and Levelt (1995) used a technique where participants monitored their internal speech for a certain phoneme that could occur at different positions in the word. Participants heard an English word and had to silently generate the Dutch translation of that words, while they were also monitoring their internal production for a target phoneme which had been (auditorily) presented to them before. In bisyllabic target words (e.g., *deken* "blanket"), participants responded faster to the onset of the first syllable than to the onset of the second syllable. For instance, they responded faster to the /d/ than to the /k/ in *deken*. This shows that encoding the first syllable starts before the encoding of the second syllable. In a syllable monitoring task, participants detected syllables faster when those corresponded to the target's first syllable, like /ma:/ in *mager* "thin" and /ma:x/ in *maagden* "servants" - a similar effect like the effect found in comprehension by Zwitserlood et al. (1993) for Dutch. This finding suggests that participants monitor a syllabified phonological representation. In a final experiment, participants monitored for all four consonants of a CVC.CVC-word like *lifter* ("hitch-hiker"). Monitoring latencies increased with the position of the target in the word, again suggesting the left-to-right direction of phonological encoding.

Sevald and Dell (1994) investigated the planning of CVC-sequences in a paradigm where participants had to produce a sequence of CVC-items as many times as possible within eight seconds. The CVCs could be of various combinations, including, for instance, a complete repetition condition (e.g., *PICK PICK PICK PICK*), and repetitions of the rhyme (*PICK TICK PICK TICK*), or the CV (*PICK PIN PICK PIN*). The repetitions could moreover be alternating (= 'far repetition': *PICK PIN PICK PIN*) or follow the pattern *PICK PIN PIN PICK* (= 'near repetition'), and this was again crossed with the onset, nucleus and coda, which could follow different repetition patterns. Participants managed to produce the highest number of CVCs when those were CVC-repetitions. Repeating the coda consonant helped, while repetition of the onset inhibited. Moreover, no facilitative effect for repeating the rhyme was found. These results again suggest that segments within a syllable are not produced simultaneously, but sequentially from left to right.

In sum, in contrast to the speech error data, the experiments reveal evidence for the time course of phonological encoding. They suggest that the segments within a syllable become available from left to right. In multisyllabic words, syllables become available sequentially from left to right, too.

Levelt's Model of Phonological Encoding

Levelt's model of phonological encoding (Levelt, 1992, 1993) is the basis for the experiments reported in this thesis. The precursors of this model, like Shattuck-Hufnagel's (1979, 1983) symbolic slot-and-filler-theory or Dell's (1986) interactive-activation model, were inspired by speech error data. Segments and syllable frames become available independently, and the segments then associate to positions in the frame (for a discussion of these models see, e.g., Levelt, 1989; Meyer, 1992). The slightly extended and modified later versions of the models (Dell, 1988, Shattuck-Hufnagel, 1992) reflected the reinterpretation of the syllable-position constraint in speech errors as a word-onset constraint. Crucially, however, these models had been designed for the encoding of single words. As was noticed already by Shattuck-Hufnagel (1979), there remains the question of why speakers should produce segments and slots in two independent processes at all. In Shattuck-Hufnagel's and Dell's models, this make-up of the phonological encoding mechanism has no other functionality than to explain why certain speech errors occur.

However, we normally produce not only single words, but strings of connected speech. In connected speech, lexical word boundaries do not necessarily align with syllable boundaries: When we say *she gave it*, the syllables we produce will be $(\text{ʃɪ:})_\sigma(\text{gɛɪ})_\sigma(\text{vɪt})_\sigma$ - the last segment of the word *gave* [v] will surface in the onset of the following syllable. The same holds for the Dutch encliticized forms: *koop het* "buy it" syllabifies $(\text{ko:})_\sigma(\text{pət})_\sigma$. Encliticization is frequent in spontaneous speech, which is the kind of output a speaker normally produces. The fact that the surface syllable structure does not always coincide with the syllables of the single lexical words poses problems for models of phonological encoding that specify syllabic positions in frames and label segments for syllabic positions. Neither Shattuck-Hufnagel's nor Dell's model could explain why the stop in the coda of *koop* surfaces in onset position of the second syllable $(\text{pət})_\sigma$ in the Dutch example mentioned above. In both models, the stop is specified as a coda segment of the verb *koop* that has to go into a coda position in the syllable or word frame.

Levelt (1992, 1993) suggested that the separation of segments and frames to which the segments are then associated receives a function when we want to account for the generation of connected speech, especially the encliticized forms. From a processing point of view, it does not seem to be useful to first construct fully syllabified forms for individual lexical items (like for *koop* and *het*) that never appear in the output of the production process. Instead, one would prefer to produce the postlexical surface syllables immediately. Levelt therefore proposed that a speaker creates frames of the size of a prosodic word as the basis for syllabification instead of single frames for individual lexical units, see Figure 2.

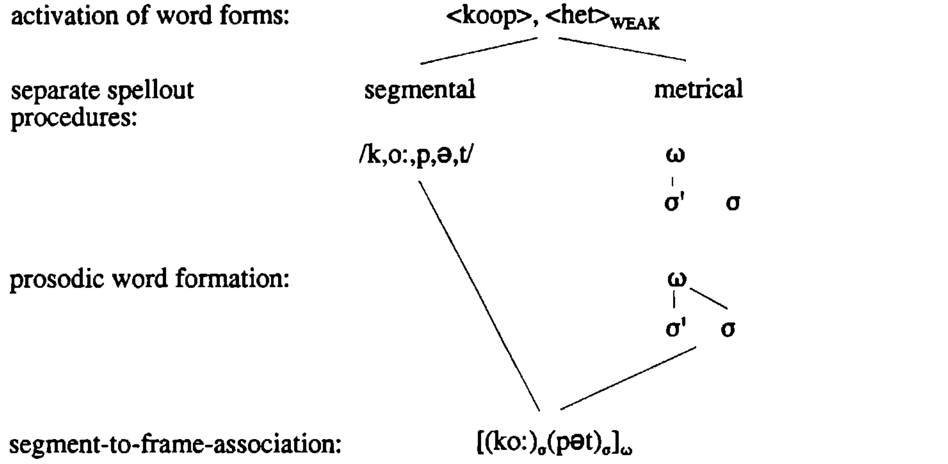


Figure 2. The Phonological Encoding Component of Levelt's Model of Language Production

Two separate procedures, a segmental spellout procedure and a metrical spellout procedure, do not work on each single lexical unit, but on lexical words that together build a prosodic word (e.g., the verb *koop* and the clitic *het*). The segmental spellout procedure delivers an ordered sequence of segments (e.g. /k,o:,p,ə,t/). Crucially, and in contrast to the models mentioned above, these segments are not marked for syllable positions. The second procedure in phonological encoding, metrical spellout, makes available information on the number of syllables a word consists of, and its stress pattern (see also Meijer, 1994).

In the process of prosodic word formation, the metrical information of the single lexical words (e.g., 'monosyllabic prosodic word' for *koop*, and 'monosyllabic unit, no prosodic word' for the enclitic *het*) is integrated into one metrical frame for the whole prosodic word, resulting in a structure like $(\sigma' \sigma)_\omega$. This frame is then combined with the sequence of segments in a process of segment-to-frame-association, the result of which is the surface syllable structure. Crucially, since the segments are not specified for syllable positions, the association process will follow the general phonological rules for syllabification like the rule that vowels constitute syllable nuclei, the decrease of sonority from syllable nuclei to syllable edges, and the maximal onset principle (cf. Levelt, 1992, p. 14f). Producing separately an ordered string of segments not specified for syllable positions and frames for prosodic instead of lexical words has the advantage that only those syllables are produced that surface in speech output. Thus, the syllables $(\text{ko:})_\sigma$ and $(\text{pət})_\sigma$ are produced, but not the syllables $(\text{ko:p})_\sigma$ and $(\text{ət})_\sigma$, that correspond to the lexical units, but do never surface. This is economical and elegant from a processing point of view.

A further assumption of Levelt's model is that the phonological syllabic units are used to access a syllabary to phonetically encode the utterance (Levelt, 1992, 1993; Levelt & Wheeldon, 1994). In the syllabary, the phonological syllables that evolve from the syllable-to-frame-association process are paired with a stored abstract gestural score for that syllable. Storage of syllable motor programs had first been suggested by Crompton (1982). The gestural scores in the syllabary are still assumed to be abstract and specify the articulatory tasks that have to be performed by the different articulators, not the specific motor programs that have to be executed to realize these tasks. It is likely that not all syllables are stored, but only those that are sufficiently frequent to have become 'overlearned'. Evidence for the syllabary comes from a study by Levelt and Wheeldon (1994). The authors based their argument on the frequency effect. When words are retrieved from the lexicon, high-frequent word forms are produced faster than low-frequent ones (e.g., Balota & Chumbley, 1984, 1985; Jescheniak & Levelt, 1994). Similarly, if later syllables have to be retrieved from a syllable store - the syllabary - another frequency effect should occur, this time of syllable frequency. Words made from low-frequent syllables should be produced more slowly than words made from high-frequent ones, and this effect should be independent from the frequency of the whole word.

Levelt and Wheeldon varied word and syllable frequency independently in a task where participants learned to associate a sequence of abstract symbols (like *[[[[[*) with a certain word and had to name the word as soon as the symbols appeared on a screen. In one experiment, participants produced high- and low-frequent words that consisted of either two high-frequent syllables or two low-frequent ones. Indeed, two independent frequency effects were obtained, one or word frequency - high-frequent words were named faster than low-frequent ones - and one for syllable frequency - words made from high frequent syllables were named faster than words made from low frequent syllables. A potential problem with this experiment is the fact that the syllables preceding low-frequent second syllables were mostly CVC- or CVV-syllables, of which the latter all contained diphthongs. The syllables preceding the high frequent syllables, on the other hand, were mostly CVV-syllables with long vowels (not diphthongs) and CVs. Thus, low-frequent syllables were preceded by a syllable with a more complex gestural score and this might have influenced the naming latencies. This problem was avoided in a further experiment, which showed that the frequency of the second syllable affected naming latencies when the frequency of the first syllable and the frequency of the whole word were held constant. The syllabary account nicely explains the finding of a syllable effect independent from word frequency. Moreover, the authors found in a control experiment that the number of phonemes in a syllable did not influence the results. This is again predicted by the syllabary account, since gestural scores are retrieved from the syllabary as whole units. The importance of the second syllable is accounted for by the assumption that a speaker will only start articulating a word when its phonetic encoding is complete (see also Levelt, 1989). One might want to test whether this still holds for longer prosodic words like *bedelares* ("female beggar"), or *koop 'k ens een* ("buy I once a"). As the authors note themselves, their materials had been controlled for word and syllable frequency, but not for the frequency of demisyllables, and also not for phoneme frequency. If high-frequent syllables turned out to be composed of high-frequent subsyllabic units, the point for the syllabary would be weakened. More experiments are needed to prove the existence of the syllabary.

Predictions of Levelt's Model

With respect to syllabification in connected speech, Levelt's model makes two strong and related predictions:

First, there is an important prediction regarding the time course of phonological encoding: Syllabification is a late process. Syllables are created only after the phonological segments have been spelled out and the prosodic frame has been created. At an earlier point in time, the segments may be available, but syllables should not yet play a role. This prediction was tested in the experiments reported in Part 1 of this thesis.

The second prediction concerns the number of levels of syllable structure which are involved in phonological encoding. One could assume that a speaker first produces syllables for the individual words, for instance, *koop* $(\text{ko:p})_\sigma$ and *het*, $(\text{ət})_\sigma$. In a next step, these syllables are concatenated, and the consonant preceding a vowel-initial syllable will resyllabify into the onset of this syllable, like the /p/ in the encliticized form *koop het* $(\text{ko:})_\sigma(\text{pət})_\sigma$. This involves two levels of syllable structure: The underlying and abstract lexical syllables of the individual words, and the surface syllables. In contrast, Levelt's model of phonological encoding does not assume an intermediate level of lexical syllables. Instead, speakers immediately produce surface syllables. The model hence predicts that these surface syllables are the only syllables that are constructed during phonological encoding. This was tested in the experiments in the second part of this thesis.

PART 1: THE TIME COURSE OF SYLLABIFICATION

A series of auditory syllable priming experiments is reported that investigated the time course of syllabification in single words and in encliticized forms. All experiments included the same kind of stimuli and targets. They aimed to find evidence for an early segmental overlap effect and a late syllable match effect as they are predicted by Levelt's model of phonological encoding.

Auditory Syllable Priming in a Paired-Associate Learning Task: Paradigm, Stimuli, and Hypotheses

The Picture-Word Interference Paradigm

As research has shown in the past decade, the picture-word interference paradigm can be used to explore the time course of language production processes. The paradigm is based on the Stroop-task: Participants have no problem in reading aloud color terms that are printed in incongruent ink colors, for instance, the word *blue*printed in red ink. They are as good at this as at reading the color terms printed in black. However, participants have great difficulties in naming the colors of incongruent words, for instance, saying "red" when the word *blue* is written in red ink. This takes much longer than naming the color of a color block or a row of colored symbols. The classic Stroop-task has been varied in several ways. An important variant in the development of the picture-word interference paradigm was when Dyer (1971) varied the stimulus onset asynchrony (SOA) by separating words and colors in time. He presented the incongruent color word in black ink at different points in time with respect to the onset of the presentation of the colored word of which participants had to name either the color or the name. The onset of the black word and the colored word were either aligned, i.e., the SOA was 0 ms, or the word preceded the color (= negative SOA) up to 500 ms. Another important step was when, also in the seventies, pictures were first used instead of colors. Incongruent words that appear as interfering stimuli (IS) with line drawings of pictures slow down participants' responses in picture naming, but incongruent pictures do not affect word naming (see, for instance, the first experiment in Glaser and Düngelhoff, 1984). See MacLeod (1991) for an overview on research on the Stroop effect.

Different kinds of IS have different effects on picture naming. Graphemically related IS facilitate picture naming. For instance, Rayner and Posnansky (1978) obtained faster RTs for nonword IS that were graphemically related to the picture, for instance, the stimulus *hcnre* for the picture of a horse, when compared to a control condition, where no IS appeared. Tousman and Inhoff (1992) in a word naming task obtained priming of a bisyllabic word like *glucose* by a visually presented prime that corresponded to the first syllable, compared to a neutral prime. Starreveld and la Heij (1995) found for Dutch that words that were orthographically related to a target, like the stimulus *kap* ("cap") for the picture of a cat (Dutch *kat*), facilitated responses compared to an unrelated baseline.

IS that are identical to the target word normally also cause facilitation. They did so in Rayner and Posnansky's study, for instance, and Glaser and Düngelhoff (1984) obtained the fastest RTs for identical IS (= "concept congruent" condition in their experiments 1 and 2). Identical IS were the only IS in this study that facilitated responses compared to a neutral baseline, which was a row of 6 Xes. Unrelated words normally inhibit the response compared to a neutral baseline (Rayner & Posnansky, 1978, Glaser & Düngelhoff, 1984). Finally, when IS are semantically related, they cause interference. Underwood (1976) reported a prolonged picture naming latency compared to an unrelated, a nonword, or a silence condition when a semantic associate occurred to the right of a picture, e.g., the word *sail* for the picture of a boat. Glaser and Düngelhoff (1984) systematically varied the timing of the presentation of the pictures and the semantically related or unrelated IS that were printed inside a line drawing of the picture. They found inhibition for unrelated IS (e.g., the word *car* superimposed on the picture of a house) and also for semantically related IS (e.g., the word *house* on the picture of a church). Semantically related IS inhibited the response to a greater extent than unrelated IS. This difference in inhibition was significant for SOAs between -100 and 100 ms.

While the earlier studies used exclusively written words as IS (for a summary see also Schriefers, Meyer, & Levelt, 1990), more recently auditory IS yielded results that are similar to the ones for visual stimuli. First, an auditory neutral condition had to be found to replace the row of Xes used in many visual studies. Schriefers, Meyer, and Levelt (1990) included an unrelated IS condition, a silence condition, where the picture was presented without IS, a blanco condition, where the IS was the spoken word *blanko* "blank", and two noise conditions. In one of them, a stretch of white noise was matched in length to the unrelated IS that went with a particular picture, in the other, the noise was 200 ms longer than the unrelated IS. It turned out that at SOA 300 ms, all IS conditions did not differ much. At the earlier SOA 0 ms, however, RTs obtained in the noise conditions and the silence condition did not differ, but the RTs for the unrelated and also for the blanco condition (though less pronounced) were significantly longer, indicating interference of the auditory stimulus with picture naming.

In auditory picture-word interference studies, IS exhibit various relations to the pictures' names which affect response latencies. Schriefers et al. found that semantically related IS, e.g., *nijlpaard* ("hippopotamus") for a picture of a crocodile (Dutch *krokodil*), showed significantly more inhibition than unrelated IS at an early SOA of -150 ms, when compared to a silence condition. On the other hand, phonologically related primes, e.g., *krokus* ("crocus") for the crocodile picture, showed significantly less inhibition at SOA 0 ms and even facilitated responses relative to the silence baseline at SOA 150 ms, again differing significantly from unrelated IS. This can be taken as evidence for the fact that the semantic processing of a target starts before it is phonologically encoded, as most models of speech production assume. In addition, a blanco condition inhibited responses to a lesser degree than the unrelated condition, but this difference was not significant anymore at SOA 150 ms. As mentioned already in the introduction, Meyer and Schriefers (1991) obtained facilitation for IS that were phonologically related to the begin of targets at SOAs -150, 0, and 150 ms, while end-related targets showed facilitation only at SOAs 0 and 150 ms.

The production experiments so far used whole words (e.g., target: *krokodil*, phonologically related IS: *krokus*). In the experiments that I will present in the first part of this thesis, syllables occurred as auditory IS instead of whole words. As a neutral IS condition, a stretch of pink noise was presented, which was matched in length to the mean length of all IS in a specific experiment. As compared to white noise, which has a uniform distribution of energy over the whole spectrum, pink noise has less energy in the higher frequency part and is therefore closer to the speech signal than white noise. Two arguments speak against a blanco condition as a neutral baseline. First, the pattern over the SOAs that Schriefers et al. (1990) had obtained was difficult to interpret, and more importantly, the "neutral" word *blanco* is nevertheless phonologically related to some target words with which it shares one or more segments, or the stress pattern, the number of syllables, and so on. The reason to choose pink noise and not a silence condition was that the trials in the neutral and the experimental conditions are more similar: Participants hear IS on their headphones in all trials.

The Semantic-Associate Learning Task

Experimental research in language production always faces the problem of finding a paradigm that on the one hand elicits exactly the utterance the experimenter wants the participant to produce, and on the other hand is still comparable to a natural situation of language production. Since the materials used in the experiments could not be depicted, a different technique than picture-naming had to be used to elicit the target words. Several methods have been reported. La

Heij et al. (1990) used a translation task. Participants had to translate an English word into its Dutch equivalent. This task has been used in several studies recently, like Jescheniak and Levelt (1994), Meijer (1994), Wheeldon and Levelt (1995). La Heij, Starreveld, and Steehouwer (1993) used a definition naming task, where participants for instance had to respond *green* when the target was *The color of lettuce*, or *sheep* to *An animal that bleats* (see also Starreveld & La Heij, 1995). Participants in a study by Levelt and Wheeldon (1994) learned pairs of target words and abstract symbols, and produced the target word when the corresponding symbol (e.g., *]]]]]]*) appeared on the screen.

I opted for a semantic-associate learning task (Meyer 1988, 1990, 1991). In this task, participants receive a sheet of paper with a list containing pairs of words that are semantically associated (e.g., *jacket - clothing*). They are instructed to learn these pairs by heart, in such a way that they are able to produce the second member of a pair as soon as the first member occurs on a screen. Compared to definition naming, this task has the advantage of short eliciting stimuli. Participants do not have to process a whole sentence, but only a word as the cue for the target utterance. In contrast to the translation task, only one language is involved, and additional factors due to bilingual processing are avoided.

In my experiments, the semantically associated pairs consisted of a noun and a verb, for example, the noun *eten* "meal" and the verb *koken* "to cook". The nouns were used as visual cues in written form to elicit the production of the verb forms. The orthographic presentation allowed abstract nouns like *succes* ("success") as well as concrete nouns to be included. With respect to the targets, this task allowed for a certain degree of variation. Participants could not only be instructed to produce the infinitive forms of the verbs, but also past tense forms like *kookte* ("cooked") or the verb with an encliticized pronoun, like in *cook het* ("cooked it") or *roep hem* ("call him"). This was a further advantage compared to a picture-naming or a translation task.

The Targets

In all but one experiment, the targets were verbs. These could occur either in the <u>infinitive</u>, like *koken* ("to cook"), or in the <u>past tense</u> form like *kookte* ("cooked, sg."), or in an <u>encliticized</u> form, where the verb was followed by the schwa-initial weak form of the function word *het* ("it") as in *kook het*. Importantly, while the sequence of initial segments is the same for all three forms, /ko:k/ in this case, the forms differ in the location of the syllable boundary: In the infinitive and the encliticized form, the target's first syllable is one segment shorter than the first syllable of the past tense verb, compare *koken* $(\text{ko:})_\sigma(\text{kən})_\sigma$ or *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ with *kookte* $(\text{ko:k})_\sigma(\text{tə})_\sigma$.

The IS Conditions

The IS were of five different kinds: They were either phonologically related or phonologically unrelated to the target word. In the related IS conditions, they corresponded to the target word's first phonemes, like for instance /ko:/ or /ko:k/ for *koken*, and *kookte*. Unrelated IS were taken from a different target word from the stimuli set, like /le:/ or /le:r/ from *leren* $(\text{le:})_\sigma(\text{rən})_\sigma$ ("to teach"). The stem of this word had no segments in common with the target's stem. On the basis of the experiments that had been run so far, participants should react faster when they hear a phonologically related IS and more slowly when they hear an unrelated one.

Targets:			
Infinitive:	*koken*	$(\text{ko:})_\sigma(\text{kən})_\sigma$	"to cook"
Encliticized:	*kook het*	$(\text{ko:})_\sigma(\text{kət})_\sigma$	"cook it"
Past Tense:	*kookte*	$(\text{ko:k})_\sigma(\text{tə})_\sigma$	"cooked"
Interfering Stimuli:			
	Short	Long	
Phonologically Related:	ko:	ko:k	
Phonologically Unrelated:	le:	le:r	
Pink Noise			

Table 2. Example for Targets and IS Conditions in Experiments 1 to 5

Furthermore, the length of the IS was varied. IS were either short (like /ko:/ and /le:/) or long (like /ko:k/ and /le:r/). A general length effect is likely to occur, since the longer the interfering signal, the more slowly participants will probably react. Long IS should thus yield longer RTs than short ones. This should hold for related as well as for unrelated IS.

To test whether syllable structure has an effect, the IS either did or did not correspond to the target's first syllable. For targets in the infinitive or encliticized forms, the short IS corresponded to the first syllable. For targets in the past tense form, the long IS corresponded to the first syllable. To test whether phonologically related IS that match the first syllable provide more priming than phonologically related IS that do not match the first syllable, I looked at the gain from related short IS compared to the gain from related long IS. To measure the amount of priming that a related IS provided independently from the expected general length effect, the RTs achieved with the related IS were subtracted from the RTs that were obtained with the unrelated IS of corresponding length. Thus, the amount of priming provided by long related IS results from calculating the difference scores for unrelated and related long IS, and the same for the short IS alike. I will refer to this difference in priming of related and unrelated IS as the 'prime efficiency' of IS. If the difference between related and unrelated short IS exceeded the difference for related and unrelated long IS, the short IS were 'more efficient primes' than the long ones.

Importantly, the unrelated and related IS for a target were matched with respect to their CV-structure: They had the same number of slots for consonants and vowels, but filled with different segments. Thus, if the related short IS had the structure CVV, like /ko:/ in our example, the same held also for the unrelated IS, like /le:/. The two targets that were matched in CV-structure mutually provided the unrelated IS.

As a fifth IS condition, the auditory stimulus was a pink noise, matched in length to the means of all IS.

The IS were produced in different ways in different experiments. Sometimes, the IS were produced as syllables by a female speaker of Dutch, sometimes, they were cut out of the whole target verb that had been spoken by a female voice - a difference that will become important below. Table 2 shows an example for the IS conditions.

Hypotheses

Given these targets and IS conditions, the following predictions can be formulated. Firstly, based on the results of the picture-word interference experiments reported above, participants should react faster when they hear phonologically related IS than when they hear unrelated IS. How could Levelt's model account for priming of, for instance, the target word *leren* ("to teach") by the syllables /le:/ or /le:r/? Roelofs (1995, 1996) has implemented Levelt's model in a spreading-activation based network combined with a parallel system of procedures. In the network, a morpheme's abstract lemma node is linked to a form node, which points to two types of phonological information, namely the metrical structure and the segments the form consists of. The links between segment nodes and the form node are labeled for serial positions. For instance, /l/ is the first segment of *leren* and the link between /l/ and *leren* is labeled '1', while the link to the segment node /r/ is labeled '3'. Segment nodes are connected with stored syllable program nodes via links labeled for syllable constituents. For instance, the /l/ of *leren* connects to the syllable program node [le:] via an onset link, and the /r/ connects via a coda link to the syllable program node [rən]. Thus, not segments, but the links are marked for syllable positions. A consonant that can appear as a syllable onset or coda has onset and coda links.

The model fits a variety of empirical findings, among others the priming effects that have been found in Picture-Word Interference experiments. The priming effects are accounted for as follows: First, consider naming without IS. The encoding of a target form involves activation of the form node by the lemma node of, e.g. a target *leren*. The activated form node spreads activation to the metrical structure and to the segments. Since the segments of several words may be available simultaneously, a selection mechanism verifies the labels on the link which connects a node with a node one level up. The segments activate syllable program nodes. The segment node /l/ will activate not only the program node [le:], but also other syllable program nodes to which it is linked via an onset link, like [la:] or [luk]. But the node [le:] will be activated higher than the competitor syllable program nodes, because it also receives activation from the segment node /e/.

A spoken interfering word stimulus activates a cohort of words in the output form lexicon at the form node level of encoding. In addition, it also activates segments at the segmental level. Take the encoding of the monosyllabic word *leer* ("teaching"), where the form node activates the segments /l/, /e/, and /r/, which again activate the syllable program node [le:r], but also competitor syllable program nodes like [le:m], [le:k], [ke:r], or [fe:r]. Unrelated IS like *maan* ("moon") activate the form node of the members of their cohort (e.g., of *maan*, *maar*, *maat*) and the segmental nodes of the IS's segments /m/, /a/, and /n/. These in turn activate the syllable program nodes for [ma:n], and [ma:t], and also for other competitor syllable nodes like [da:t] or [la:m]. The IS thus increase the number of active competitor syllable program nodes, which in turn increases the time needed to select the target syllable program node in the network. Similarly, phonologically related IS like *leek* ("layman") will activate competitor syllable program nodes like [le:k], [le:n], or [re:k] via the segment nodes /l/, /e/, and /k/. Importantly, however, *leek* will also activate the target syllable program node [le:r]. This leads to a faster selection of this node than when unrelated IS are presented, and may even result in a facilitatory effect as compared to a neutral baseline. Among others, the model nicely fits the data obtained by Schriefers, Meyer, and Levelt (1990), and by Meyer and Schriefers (1991).

Since the spoken IS influence encoding at both the level of word form nodes and segments, the model can also account for effects of IS that are not words, but syllables, as I used in my experiments. The stimulus /le:/, for instance, will activate the cohort of nodes at form level starting with /le/ to a certain extent, and will in addition activate the segments /l/ and/e/, which in turn activate the syllable program node [le:], but also the syllable program nodes [le:n], [la:n], or [me:r]. As with the word IS, phonologically unrelated syllable IS will only activate a number of competitor syllable program nodes, which inhibit the naming response, while related IS will also increase the activation of the target syllable program node.

Secondly, in addition to the main effect of phonological relatedness, which should result in faster RTs for related than for unrelated IS, one may expect a general length effect. Due to the longer interfering signal of long IS, these should yield longer RTs than short ones. Participants should react faster to short related IS than to long related ones, and they should be faster in trials with short unrelated IS compared to trials with long unrelated IS.

Thirdly, only segmental and metrical information is available at an early stage of processing in the model of phonological encoding, but no syllables have been constructed yet. As a consequence, the target's syllable structure should play no role at that point in time. All three forms of targets - like *leren*, *leerde*, and *leer het*, or *koken, kookte, kook het* - have the same sequence of initial segments. Priming, defined as the amount of time that responses are faster for related IS as compared to unrelated IS of corresponding length, should increase when the number of segments increases that the IS overlaps with the target. Thus, long IS should be more efficient primes than short IS for all target forms alike. Or, in short: /ko:k/ should be a more efficient prime than /ko:/ for *koken*, *kookte*, and *kook het*. This effect will be called the '<u>segmental overlap effect</u>'. It should occur at an early SOA, when only segments are available, but no syllables have been built yet.

SOA variation is implemented in Roelofs' network by pre- or postponing the influence of the spoken IS for some cycles. If the spoken IS start influencing early, the activation flow in the net has just started and the IS may influence the selection of the word form node and the activation at the segmental level. Related IS will activate the target form node and the target's first segment nodes. For instance, the short related stimulus /le:/ will activate the form nodes of the cohort which includes, in addition to the target *leren*, forms like *leen*, *leven*, *leek*, and it will activate the segment nodes /l/ and /e/. The long related stimulus /le:r/ will activate the target *leren* to a higher extent than the competitor form nodes, and furthermore activate one segmental node more than short related IS do. The result is the segmental overlap effect.

A fourth effect is predicted for a later point in time, when the phonological encoder is busy with the construction of syllable structure. In contrast to the earlier stage, it should be important now that the first syllable of the verb's past tense form corresponds to the long related IS, while the first syllable of the other forms corresponds to the short IS. The model predicts what I will call a '<u>syllable match effect</u>': IS corresponding to a target's first syllable should be more efficient than IS that do not match the syllable. For infinitives and encliticized forms, the short IS should be more efficient primes than the long ones. Past tense forms, on the other hand, should be primed more efficiently by long IS than by short IS. For example, /ko:/ should be a better prime than /ko:k/ for *koken* and *kook het*, while /ko:k/

should be better than /ko:/ for *kookte*, and /le:/ should prime *leren* and *leer het* more efficiently than /le:r/, while /le:r/ should do better for the past tense verb form *leerde*.

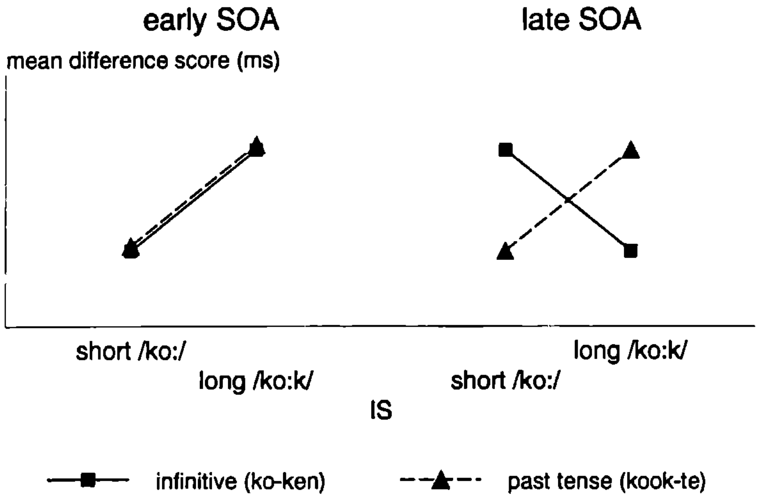


Figure 3. Pattern Predicted by Levelt's Model of Phonological Encoding for Mean Difference Scores of Unrelated and Related IS Priming the Different Target Verb Forms at an Early and a Late SOA

In Roelofs' implementation of Levelt's model, the syllable match effect is not predicted. Spoken IS affect activation at segmental level, and indirectly the selection of the syllable program nodes. Related IS like /le:/ and /le:r/ will boost the activation of the segment nodes /l/ and /e/, or /l/, /e/, and /r/, respectively, which will in turn activate syllable program nodes like [le:], [le:r], [me:r], or [ze:]. The syllable match effect could in principle be accounted for by an additional assumption. While in the current version of the network only segmental information is provided by the auditory IS, they could also carry information about the segments' syllable position. If this information is in confilct with the syllables built during the production process, one could account for the syllable match effect (Roelofs, pers. communication).

According to the strictest formulation of the predictions of Levelt's nodel, the segmental overlap and syllable match effect should occur at different SOAs. Technically speaking, an interaction is predicted of SOA, verb form (infinitive and past tense), phonological relatedness (related and unrelated), and IS length (short and long) for experiments testing single verbs like *koken* or *kookte*. The interaction of verb form, relatedness, and length should be significant at a late SOA, but not at an early one. At a late SOA, the priming obtained for short and long related IS when compared to their unrelated counterparts of equal length should depend on the target's syllable structure, which is different for the two verb forms. At the early SOA, long IS should be more efficient primes for both verb forms. For the experiment with encliticized forms like *kook het*, the model predicts a significant interaction of SOA, relatedness, and length. While long IS should be the more efficient prime at an early SOA, short IS should do better at a late SOA, because the first surface syllable is short: $(\text{ko:})_\sigma(\text{kət})_\sigma$. Figure 3 schematically illustrates the interactions.

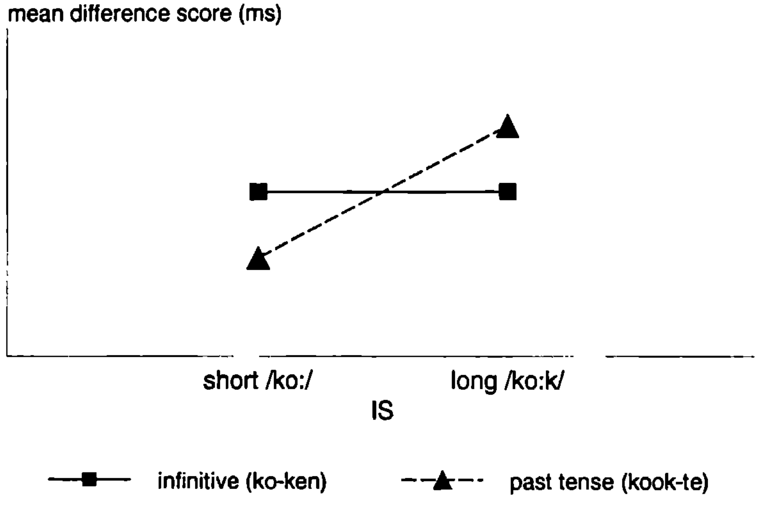


Figure 4. Pattern Predicted by the Model of Phonological Encoding for Mean Difference Scores of Unrelated and Related IS Priming the Different Target Verb Forms when Segmental Overlap and Syllable Match Effect Co-occur

However, it may as well be the case that at a certain SOA both effects co-occur, that is, that the segmental overlap is still effective and syllable match is already important. This could happen, for instance, when participants or items differ in processing speed. Under this less strict view, and under the assumption that both effects are additive, the following pattern is expected. Long IS should be more efficient than short ones in all target forms due to the segmental overlap effect. In addition, long IS should also be the more efficient primes for the past tense forms because of the syllable match effect. This should further increase their prime efficiency compared to the short IS.

In the infinitives and encliticized forms, on the other hand, long IS should be the more efficient primes due to the segmental overlap effect, while short IS should surpass long IS in prime efficiency due to the syllable match effect. As a result, the prime efficiency of short and long IS should differ less or not at all, dependent on the strength of the two opposite effects. Figure 4 schematically shows the pattern predicted under the assumption that the effects that co-occur are equally strong and additive.

Overview of the Syllable Priming Experiments

The targets in **Experiment 1** were single verbs like *koken, kookte* ("to cook", "cooked"). While the proportion of related IS was low in the first experiment, **Experiment 2** included an increased proportion of 40% related IS. Both experiments used IS that had been spoken as syllables. **Experiment 3** was a replication of Experiment 2, but in contrast to that experiment, the IS were spliced out of the target verb forms and thus closer to the phonetic level. **Experiment 4** tested encliticized forms (e.g., *kook het* "cook it") and included again a low proportion of related IS. This proportion was increased to 40% in **Experiments 5a,b**, which also tested encliticized targets. The encliticized target forms in Experiment 5a were similar to Experiment 4, while Experiment 5b tested encliticized forms of the type *kookt het*, which - as the past tense forms - have a long first syllable. **Experiment 6** and **Experiment 7** were control studies to exclude the possibility that the effects obtained were due to morphological or lexical rather than phonological factors.

It turned out that participants made more errors when they heard phonologically unrelated IS than when the IS were related. Analyses of Variance (Anovas) showed that this difference was significant in all experiments, either in the Anovas over subjects, or in the Anovas over items, or in both Anovas. But the unrelated IS, which showed the higher error rates, also yielded slower response latencies than the related IS. There was hence no speed accuracy tradeoff.

The first experiment will be described in more detail. For the remaining experiments, I will only describe what differed from the first study. Appendix A7.1 contains a table showing some characteristics of Experiments 1 to 7.

Experiment 1: Single Verbs, 10% Related IS

One aim of the first experiment was to test the experimental paradigm. It had to be tested whether spoken syllable IS yield any phonological effect at all, because in previous picture-word interference studies the auditory IS had always been whole words. Furthermore, the semantic-associate learning task had never been used with interfering IS before. Moreover, the task had been used to elicit single word responses (Meyer, 1988, 1990, 1991), but not to elicit targets containing more than one word like encliticized forms. In the first experiment, the targets were single words, because for these it was sure that the semantic-associate learning task worked. If the experiment failed to show any effects, the reason would be the auditory syllable IS rather than the eliciting task.

Auditory stimuli were presented at two SOAs, either simultaneously with the presentation of the noun on the screen (= SOA 0 ms) or 150 ms after that (= SOA 150 ms), because former studies had obtained phonological effects for these SOAs.

For many of the experiments presented below, a pretest was run to test the stimuli. When an experiment comprised a pretest, this will be reported first in the Method section, preceding the description of the main experiment.

Method

Pretest of the Stimuli

Participants. Thirty-three students participated in the pretest. They were undergraduates from the University of Nijmegen recruited from the participant pool of the Max Planck Institute for Psycholinguistics and native speakers of Dutch. They were paid HFL 8.50 for their participation. Each student took part in only one experiment or one pretest.

Stimuli. The stimuli were 33 pairs of nouns and semantically associated verbs (see Appendix A1.1). All verbs had monosyllabic stems and regular past tense forms. Their first syllable contained either a long vowel or ended in a consonantal cluster (or both) to avoid ambisyllabic consonants. The noun and the verb were phonologically unrelated. Their semantic associatedness was tested by four speakers of Dutch who had to exclude those pairs from a list that they considered not to be closely semantically associated. Only those pairs were considered for the pretest that all speakers judged to be associated.

Apparatus. The experiment was run with the NESU-setup, which has been developed at the Max Planck Institute for Psycholinguistics. Participants saw the visual stimuli in white capital letters (Arial, 21 point) centered on a black background on a high resolution NECMultisync II CRT monitor. A Hermac AT computer controlled the presentation of the visual stimuli and the collection of the response data. Participants' responses were recorded on one channel of a DAT-tape with a SONY 55 ES DAT recorder, using the microphone of Sennheiser HMD 224 headphones. On the second channel of the tape, two beeps of different frequency marked the onset of the visual cue and the triggering of the automatic registration of the RT by a voicekey.

Procedure. Students took part in single sessions. They received a sheet of paper with instructions and a separate sheet with the nouns and the semantically associated verbs in infinitive form printed on it. They learned the pairs by heart until they said they knew the pairs. This normally took five to eight minutes. The experimentor collected the sheets. Participants were told that they participated in a memorization experiment and were instructed to produce the verb as soon as the noun appeared on the screen. In the experiment, every participant completed a list that contained three blocks, within which each target occurred once. The randomization was different for every block and every participant to control for potential long term semantic priming effects.

A single trial looked as follows. The noun appeared on the screen for 500 ms. RTs were measured from the onset of the visual presentation. If there had not been a response within 2000 ms, the response was automatically coded as a missing value. After 1000 ms pause, the following trial began. A single trial lasted 3000 ms, the whole experiment took about 20 minutes.

Data analysis. Some of the responses had to be marked as missing values. Firstly, these were cases where the participants had made noises that triggered the voice key before the real response started, and secondly hesitations (filled pauses or stuttering responses). Thirdly, wrong responses were counted as errors. Fourthly, too slow or simply no responses to a target were excluded, as well as single sounds or syllables. The missing values were detected by listening to the taped recordings of the experimental session. RTs below 400 ms and exceeding 2000 ms were automatically marked as missing values. All missing values were excluded from the analysis. Eight participants were excluded because of technical errors or a high error rate.

Results of the pretest. Only the second and third repetition of an item were included in the analysis. Each item occurred 50 times (two repetitions for 25 participants). Only those pairs were chosen as experimental pairs that showed a small number of errors (lower/equal 10%), a Standard Deviation below 270 ms, and mean RTs lower than 1000 ms. The three criteria tended to co-occur (see again Appendix A1.1).

A further demand which influenced the selection of the experimental materials was that the verbs had to build pairs that mutually provided the unrelated IS. These pairs shared the number of consonants in onset position and preceding the schwa of the infinitive ending /-ən/. The final list contained 20 noun-verb pairs. Eight verbs had a CVVC-stem (e.g. *roken* "to smoke"), six had a CVCC- (e.g. *verven* "to paint"), and six a CCVC-stem (e.g. *knielen* "to kneel").

Main Experiment

Participants. Sixty-six students participated in the experiment, 35 for SOA 0 ms and 31 for SOA 150 ms.

Stimuli. The stimuli were pairs of nouns and semantically associated verbs, e.g. *soep* ("soup") - *koken* ("to cook"). Participants had to name the verbs in the infinitive (*koken*) or, in a different experimental block, in the past tense form (*kookte*).

The IS were of five different kinds: phonologically related short and long IS, unrelated short and long IS, and pink noise. The syllable IS were spoken as syllables by a female speaker who alternatingly pronounced a target verb and a syllable provided by that verb, like *koken - ko - koken - kook - leren - lee - leren - leer*, and so forth.[2] The syllable IS were recorded with a SONY 55 ES DAT recorder and a Sennheiser ME 80 microphone. They were digitized with a sample frequency of 20 kHz and the speech processing program XWAVES.

In addition to the experimental targets, there were filler words. Similar to the targets were the fillers verbs with monosyllabic verb stems. They were combined with four unrelated IS which were nonword syllables of Dutch. Of each syllable, there was a long and a short version that differed in one segment, like *prui* and *pruit*. Each version of a syllable occurred twice, combined with different filler words. In addition, fillers were combined with pink noise. As a consequence, there were more unrelated than related IS in the experiment.

[2] One could argue that the syllables had better been included in lists like *koken - ko - kookte - kook*, in which the short and long syllables both follow a word of which they are the first syllable. The preceding verb form could influence the acoustic quality of the syllable stimuli in some way. The fact that both short and long syllables were preceded by an infinitive could have made the short stimulus more similar to the infinitive's first syllable than the long stimulus is similar to the first syllable of a past tense form. Consequently, syllable match effects should occur for the infinitives but not for the past tense forms. However, in the experiments presented below, syllable match effects were obtained for past tense forms, not for infinitives.

Meyer and Schriefers (1991) raised the possibility that a high proportion of related IS may lead participants to use the IS strategically. Participants could either try to predict upcoming items on the basis of the IS when they are presented at an early SOA. Or they could check the correspondence between target and IS when the IS are presented at a later SOA. A lower proportion of phonologically related IS should discourage them from doing so. This had been the argument for Meyer and Schriefers to replicate an experiment that had been run with 50% phonologically related IS, this time with a lower proportion of 22% related IS. The manipulation did not affect the results. However, the appearance of phonologically related IS at almost a quarter of the trials might still have been too salient to make participants ignore the IS, and perhaps they used strategies in both experiments. The topic of strategic effects will return in more detail with Experiment 2.

To avoid strategic processing, the proportion of related IS was low in Experiment 1, which contained 10% related and 90% unrelated IS (not counting noise). To delimit the duration of the experiment, only 10 target verbs could be tested. The 10 experimental items occurred in two phonologically related conditions (short and long). Consequently, 180 unrelated trials had to be included. Twenty of the unrelated trials were already provided by the unrelated short and long condition of the 10 experimental items. The remaining 160 unrelated trials were provided by 40 filler items.

Apparatus. The apparatus was the same as in the pretest. Participants heard on their headphones the IS that came from the speechserver of the NESU-setup.

Design. The design comprised five types of IS (related short, related long, unrelated short, unrelated long, pink noise), and two types of verb form (infinitive, past tense). Two SOAs (0 ms, 150 ms) were tested between subjects.

Procedure. Students took part in single or double sessions. The semantic associates of 10 experimental and 40 filler pairs were too many to be learned by heart at once. Therefore, the materials were split into two sets, which were conducted in the experiment one after the other. Each set contained 5 target items and 20 fillers (see Appendix A1.2). The verb was printed on the sheet of paper in the form that had to be produced by the participant in the following experimental block - either the infinitive form (e.g. *soep* - *koken* "soup - cook"), or the past tense form (*soep* - *kookte* "soup - cooked"). Participants were told that the memorization experiment consisted of a practice and an experimental block, which were interrupted by a short pause in which more instructions could be given. Furthermore they were told that they would get four sheets of paper to learn during the experiment.

Half of the participants started with set 1, half of them started with set 2. Within each of these two participant groups, half of the participants performed the infinitive condition first, followed by the past tense condition, and half of the participants performed the verb forms in opposite order. In total, each participant completed thus four experimental lists.

For every list, participants first received a separate practice list that started with the word *Attentie* ("attention") appearing on the screen for 3000 ms. After 1500 ms pause, the practice list started. It contained two parts within which each of the 25 target and filler pairs occurred once. They were accompanied by a pink noise IS that was presented at the appropriate SOA (0 ms or 150 ms). The randomization differed in the two parts, but was the same for all participants. The practice list ended with the word *Pauze* ("pause"), which appeared on the screen for 3000 ms. In case the participant missed a pair, the experimenter had a chance to mention the correct pair. The first experimental list began, again preceded by an attention sign, with six filler items that gave participants the chance to get used to the presentation of the auditory syllable IS. The targets and fillers of the set followed, randomized in five blocks. Each block contained the five experimental targets and twenty fillers once, in different IS conditions. The first block contained the first experimental target with the related short IS, the second experimental target with the related long IS, the third experimental target with the unrelated short IS, the fourth experimental target with the unrelated long IS, and the fifth one with the pink noise, and in addition ten fillers with long and ten fillers with short IS. The second block contained the first experimental target with the pink noise, the second experimental target with the related short IS, the third with the related long, the fourth with unrelated short, and so on. Within each SOA, there was a different randomization for each experimental block and participant, except for those participants who were run in parallel, who got the same randomization. The order of the blocks was rotated between participants. Five participants started with the first block, five with the second block, and so on. No subsequent verbs in a list shared an onset or the vowel, nor were a preceding noun or IS phonologically related to the target verb. There were no more than two consecutive experimental trials. After participants had performed in the first verb form, they received the paper that showed pairs with the verbs in the other verb form. A participant produced both verb forms of a set in the same randomization.

A single trial looked as in the pretest. In addition, auditory IS were presented simultaneously with picture onset (SOA 0 ms), or 150 ms after picture onset. The whole experiment took approximately one hour.

Data analysis. Missing values were replaced by cell means. Only the experimental targets were analyzed. Two participants had to be excluded because of technical errors. As in all experiments reported, 25 participants were analyzed at each SOA. Surplus participants were randomly excluded.

Two different sets of Anovas were run. Each set contained an analysis over subjects and an analysis over items. The first set comprised as within-subjects variables 'IS' (five levels: related short, related long, unrelated short, unrelated long, pink noise), and 'verb form' (two levels: infinitive and past tense). The second set of Anovas had three within-subjects variables. These were two variables for IS, which were crossed: IS were either phonologically related or unrelated (= 'phonological relatedness' with two levels: related and unrelated), and they were either short or long (= 'length' with two levels: short and long). The third within-subjects variable was 'verb form' with two levels: infinitive and past tense. Furthermore, in both sets of Anovas the variable 'SOA' (two levels: 0 ms and 150 ms) was a between-subjects variable in the analyses over subjects, and a within-items variable in the analyses over items.

In the first set of Anovas, which comprised one variable 'IS' with five levels, the same pattern was obtained in all experiments. The effect of IS was always significant. Unrelated IS led to the highest RTs, followed by pink noise. The related IS yielded the fastest responses. The effect of SOA was significant in the analysis over items only, and no effect of verb form was obtained. When the interaction of SOA and IS was significant, which was the case in Experiments 1, 3, 4, and 5a,b, tests of simple effects revealed that the effect of IS was significant at each SOA. In the following, only the results of the second set of Anovas will be reported. Results of tests of simple effects can be found in Appendix B.

Results and Discussion

The mean RTs are shown in Table 3, see also Figure 5. Although the verb forms differed in tense, the mean RT was identical for infinitives and past tense forms (807 ms). For the infinitives, the length of related and unrelated IS did not have an effect (2 ms difference in RTs for related IS, 4 ms difference for unrelated IS). Moreover, at SOA 0 ms, the unrelated IS behaved like pink noise, while related IS led to faster RTs. At SOA 150 ms, unrelated IS yielded longer RTs than pink noise. For the past tense forms, long related IS differed from pink noise by only 10 ms at SOA 0 ms and 18 ms at SOA 150 ms, while short related IS led to the fastest RTs at both SOAs. Unrelated short and long IS again patterned alike and led to longer RTs than pink noise at both SOAs.

The effect of SOA was significant over items only ($F_2(1;9) = 13.86$, $MSE = 1940$, $p < .01$). There was a highly significant effect of phonological relatedness with $F_1(1;48) = 97.61$, $MSE = 5364$, $p < .01$ and $F_2(1;19) = 36.91$, $MSE = 5681$, $p < .01$. Participants reacted faster when they heard related than when they heard unrelated IS (mean RT related = 750 ms, unrelated = 883 ms). The interaction of SOA and relatedness was significant in the subject analysis ($F_1(1;48) = 5.65$, $MSE = 5364$, $p < .03$; $F_2(1;9) = 4.1$, $MSE = 2999$, $p < .08$). While responses in the related IS condition were 55 ms faster than in the unrelated condition at SOA 0 ms, the difference was 90 ms at SOA 150 ms. Tests of simple effects showed that this difference was significant at both SOAs. The main effect of length was not significant (mean RT for short IS = 804 ms, mean RT for long IS = 809 ms), nor was any other interaction.

In all conditions, short IS were slightly more efficient primes than long ones, but this difference was not significant, as tests of simple effects revealed. Short and long IS primed infinitives and past tense forms with similar efficiency at each SOA.

In sum, phonologically related IS led to significantly faster RTs than unrelated IS. The facilitation that had been found for phonologically related word

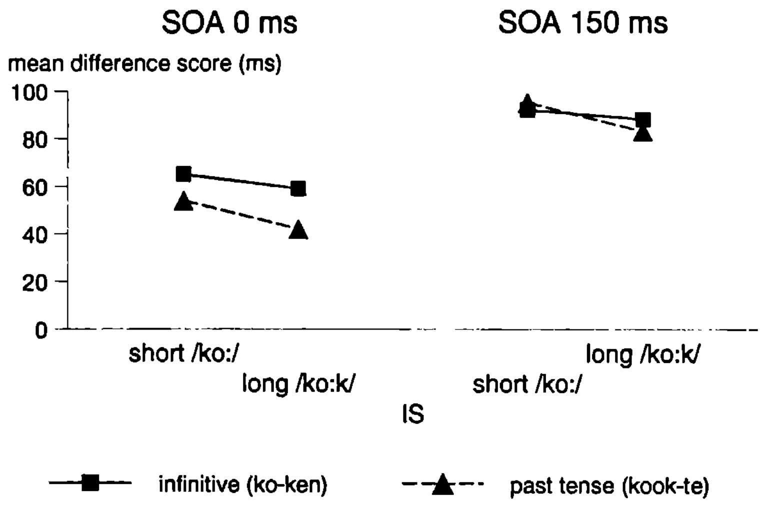


Figure 5. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Infinitive and Past Tense Forms at the Two SOAs in Experiment 1

IS in previous picture naming experiments can thus also be obtained by spoken syllables in a semantic-associate learning task. Moreover, SOAs of 0 and 150 ms seem to be the appropriate time window to tap into phonological processing.

However, neither a main effect of length was obtained, nor did one of the interactions of two or more of the variables relatedness, length, verb form, and SOA reach significance. Short and long IS were equally efficient primes for both types of target verb forms at both SOAs. This may be due to the small number of 10 items in the experiment.

	Verb Form			
IS	Infinitive		Past Tense	
	SOA 0 ms			
Short				
Related	757	(2.4)	762	(3.2)
Unrelated	822	(5.2)	816	(6.0)
Δ	65		54	
Long				
Related	759	(4.4)	787	(4.0)
Unrelated	818	(6.0)	829	(5.2)
Δ	59		42	
Pink Noise	815	(3.6)	797	(6.8)
	SOA 150 ms			
Short				
Related	778	(1.2)	766	(2.9)
Unrelated	870	(4.0)	861	(6.0)
Δ	92		95	
Long				
Related	773	(2.8)	782	(4.8)
Unrelated	861	(2.0)	865	(4.4)
Δ	88		83	
Pink Noise	797	(4.4)	804	(3.2)

Table 3. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 1
Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses.

Experiment 2: Single Verbs, 40% Related IS

Experiment 1 had only included 10 targets to allow for a low proportion of related IS in the materials. Because of the small number of targets, this experiment might have been too weak to reveal the predicted interactions. In the following experiment, the proportion of related IS was increased to 40% by adding 10 experimental targets to the 10 targets of the first experiment and not including any fillers except some practice items. The reason was that an increased proportion of related IS may also increase the chance that participants start using the IS strategically. If a syllabic match effect does exist, the fact that participants use the IS might increase the chance to pick it up.

Several strategies have been proposed that participants may use in primed lexical decision tasks in language comprehension. These can operate before a lexical entry is actually accessed, as in the creation of an expectancy set (Cañas & Bajo, 1994). Strategic processing is also possible after an entry is accessed, for instance, in a semantic matching strategy (Neely, Keefe, & Ross, 1989), or a meaning integration process for semantically related primes ('post-lexical coherence checking', de Groot, 1984, 1985), or a general checking process (Radeau, Morais, & Dewier, 1989). It has been established that changes in the proportion of related visually presented primes may induce strategies in lexical decision tasks (e.g., Tweedy & Lapinski, 1981; Tweedy, Lapinski, & Schwaneveldt, 1977; den Heyer, Brand, & Dannenberg, 1983; de Groot, 1984; Seidenberg, Waters, Sanders, & Langer, 1984) as well as in naming (Keefe & Neely, 1990). Proportion effects have also been found with phonological primes in spoken language comprehension, for instance, in perceptual identification tasks (Goldinger, Luce, Pisoni, & Marcario, 1992). In sum, two kinds of strategies have been proposed for the lexical decision task. One possibility is that participants create an expectancy set on the basis of the prime, and when the target matches one of these words the "yes / word"-response will become faster. The same has been proposed for naming (Keefe & Neely, 1990). As a second strategy, participants may check the relation between prime and target before they make a decision. This takes longer for unrelated than for related primes. In both cases, an increased proportion of related primes makes a strategic usage of the prime by the participants more likely.

Turning to the area of production, increasing the proportion of congruent trials in Stroop experiments increased the interference by incongruent trials, too. Glaser and Glaser (1982) varied in a Stroop-task the proportion of congruent trials from 30% to 80%. As a consequence, the size of facilitation and inhibition effects in the color-naming task was increased by four to five times, while their time course was not changed. The authors interpreted this as resulting from a verbal code that participants generate for the color component of a stimulus in the high proportion experiment because of its high predictive power. To conclude, an increased proportion of related IS trials is likely to increase facilitative and inhibitory effects in perception as well as in production tasks.

In my first experiment, the inclusion of even a low proportion of congruent, phonologically related IS could already have made participants attentive to the IS. Increasing the proportion of related IS might increase the obtained facilitative effects, as it did in the Glasers' study. In turn, the chance of picking up the predicted segmental overlap and syllable match effect should grow.

Experiment 2 differed from Experiment 1 in the number of targets, which was doubled to 20, and in the proportion of related IS, which was increased from 10% to 40%. The former step should provide more power, the latter step should make participants use the IS strategically.

Method

Participants. Seventy-two students participated, 42 for SOA 0 ms, and 30 for SOA 150 ms.

Stimuli. All 20 targets that had evolved from the pretest for Experiment 1 were used as stimuli for Experiment 2. There were no fillers.

Design. The design was the same as in Experiment 1.

Procedure. Since no fillers were included, participants could learn all 20 targets and six practice items at once. Half of the participants started with the infinitive condition and continued with the past tense condition, and half of them started with the past tense condition and continued with the infinitive forms. Not more than two consecutive trials contained the same IS condition and at least two trials separated two occurrences of a single item. The materials can be found in Appendix A2.1.

Data analysis. Two participants had to be excluded due to technical errors, and in addition, nine participants were excluded who made more than 10% errors. Of the remaining participants, 25 participants were randomly chosen for each SOA and the rest were excluded from the analysis.

Results

Table 4 shows the results of Experiment 2, see also Figure 6. As in Experiment 1, verb form had no effect (mean RT infinitive = 817 ms, past tense = 812 ms).

SOA was significant over items only ($F_1(1;19) = 68.97$, $MSE = 2448$, $p < .01$). As in the preceding experiment, a highly significant effect of phonological relatedness was obtained ($F_1(1;48) = 239.99$, $MSE = 7287$, $p < .01$ and $F_2(1;19) = 202.72$, $MSE = 6905$, $p < .01$). Phonologically related IS yielded faster RTs than unrelated ones (mean RT related = 750 ms, unrelated = 883 ms). The main effect of length was significant, too ($F_1(1;48) = 15.11$, $MSE = 1323$, $p < .01$; $F_2(1;19) = 14.89$, $MSE = 1089$, $p < .01$). Participants were slower when they heard long IS

	Verb Form			
IS	Infinitive		Past Tense	
	SOA 0 ms			
Short				
Related	765	(4.0)	768	(3.2)
Unrelated	902	(6.2)	885	(5.8)
Δ	137		117	
Long				
Related	783	(3.2)	765	(2.8)
Unrelated	921	(5.8)	926	(5.8)
Δ	138		161	
Pink Noise	850	(4.8)	826	(5.8)
	SOA 150 ms			
Short				
Related	729	(2.0)	730	(3.0)
Unrelated	843	(6.6)	852	(4.2)
Δ	114		122	
Long				
Related	735	(3.4)	728	(3.2)
Unrelated	863	(5.2)	868	(5.0)
Δ	128		140	
Pink Noise	782	(5.0)	773	(4.0)

Table 4. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 2
Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses.

than when they heard short ones (mean RT short = 809 ms, long = 824 ms). The interaction of relatedness and length was significant on .05-level ($F_1(1;48) = 5.65$ $MSE = 1638, p < .03$; $F_2(1;19) = 6.78, MSE = 1097, p < .02$). While RTs for short and long IS differed in 5 ms for the related IS, the difference for the unrelated IS was 24 ms, and the relatedness effect was higher for long IS (142 ms) than for short ones (123 ms). Long IS hence provided more facilitation than short IS.

Furthermore, the interaction of SOA and relatedness was significant in the analysis over items ($F_1 < 1$; $F_2\,(1;19) = 4.99, MSE = 646, p < .04$): The prime efficiency of related IS was slightly higher at SOA 0 ms (139 ms) than at SOA 150 ms (125 ms).

The interaction of verb form, relatedness, and length was not significant over subjects ($F_1(1;48) = 2.58$, $MSE = 1324$, $p < .12$), but significant over items ($F_2(1;19) = 4.97, MSE = 537, p < .04$). The difference in prime efficiency between long and short IS was higher in the past tense (31 ms) than in the infinitive (7 ms) condition. In both verb form conditions, long IS were more efficient primes than short ones.

The four-way interaction of SOA, verb form, relatedness, and length was not significant. As in Experiment 1, the interaction of verb form, relatedness, and length was investigated for each SOA separately. It was not significant at SOA 150 ms (all $F < 1$). At SOA 0 ms, however, it was significant at .05-level ($F_1(1;48) = 4.14, MSE$

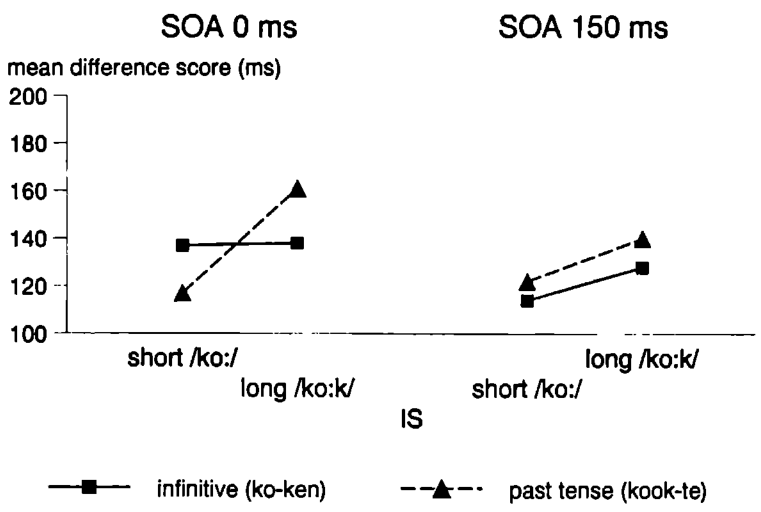


Figure 6. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Infinitive and Past Tense Forms at the Two SOAs in Experiment 2

= 1324, $p < .05$; $F_2(1;19) = 4.87$, $MSE = 888$, $p < .04$). Short IS were equally efficient as long IS in priming the infinitive verb forms at SOA 0 ms (priming effect for short IS = 138 ms, for short IS = 137 ms). However, long IS primed the past tense forms better than short IS (priming effect for long IS = 161 ms, for short IS = 117 ms). Tests of simple effects confirmed that the interaction of relatedness and length was significant for the past tense forms at SOA 0 ms, but not for the infinitives. This pattern had been predicted for the case that segmental overlap effect and syllable match effect co-occurred. While the effects cancel each other for infinitive targets, long IS should be more efficient than short ones for the past tense forms, since they provide more segmental overlap and correspond to the targets' first syllable, and this is what we find at SOA 0 ms.

For both SOAs separately, tests of simple effects were run to investigate the interaction of relatedness and verb form within each level of IS length. It turned out that the prime efficiency of short IS did not significantly differ for infinitives and for past tense forms - neither at SOA 150 ms nor at SOA 0 ms. The same held for the long IS. But importantly, the interaction of verb form, relatedness, and length was significant at SOA 0 ms, as was the interaction of relatedness by length for the past tense forms at that SOA.

Reanalysis with the Items of Experiment 1

To find out whether it was the double number of items alone that caused the significant interaction of verb form, relatedness, and length at SOA 0 ms in this experiment, a reanalysis was done. Only the 10 items were included that had also been tested in Experiment 1. If the interaction is still obtained, it is clearly the higher proportion of related IS which caused its appearance. Table 5 shows the results of the reanalysis.

Mean RT in the reanalysis was 791 ms for the infinitive forms and 784 ms for the past tense forms. The effect of phonological relatedness was highly significant with $F_1(1;48) = 160.59$, $MSE = 9176$, $p < .01$ and $F_2(1;9) = 82.52$, $MSE = 6690$, $p < .01$, responses in the related condition being faster than in the unrelated condition (mean RT related = 732 ms, unrelated = 849). The main effect of length was significant, too, but weaker in the item analysis ($F_1(1;48) = 8{,}83$, $MSE = 2692$, $p < .01$; $F_2(1;9) = 8.97$, $MSE = 1278$, $p < .02$). Long IS yielded longer RTs than short ones (mean RT short = 782 ms, long = 799 ms). The interaction of relatedness and length was significant again ($F_1(1;48) = 6.89$, $MSE = 2704$, $p < .02$; $F_2(1;9) = 7.56$, $MSE = 908$, $p < .03$). While RTs obtained with short and long IS differed in 2 ms for the related IS, the difference for the unrelated IS was 29 ms, and the relatedness effect was higher for long IS (130 ms) than for short ones (104 ms). SOA remained significant in item analysis only ($F_2(1;9) = 46.37$, $MSE = 1946$, p

< .01). The interaction of SOA and relatedness was not significant anymore.

The interaction of relatedness and length was significant within the past tense verb forms ($F_1(1;48) = 3.74$, $MSE = 2634$, $p < .06$; $F_2(1;9) = 9.31$, $MSE = 425$, $p < .02$), where the prime efficiency of long IS exceeded the efficiency for short IS by 18 ms. However, tests of simple effects showed that the interaction of relatedness and length was not significant for the past tense condition when looking

IS	Verb Form: Infinitive		Verb Form: Past Tense	
		SOA 0 ms		
Short				
Related	747	(4.0)	751	(4.4)
Unrelated	867	(4.8)	851	(6.0)
Δ	120		100	
Long				
Related	767	(2.4)	743	(3.6)
Unrelated	907	(4.8)	881	(4.8)
Δ	140		138	
Pink Noise	819	(5.2)	798	(3.6)
		SOA 150 ms		
Short				
Related	714	(1.6)	707	(3.2)
Unrelated	798	(7.2)	821	(4.0)
Δ	84		114	
Long				
Related	718	(3.2)	706	(2.0)
Unrelated	830	(6.4)	839	(3.6)
Δ	122		133	
Pink Noise	739	(5.2)	746	(2.8)

Table 5. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors for the 10 Items in Experiment 2 that had been Tested in Experiment 1
Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses.

at each SOA separately. The interaction of relatedness and length was not significant within the infinitives. The interaction of verb form, relatedness, and length was not significant either (all $F < 1$). Tests of simple effects showed that the interaction did not reach significance at any SOA (all $F < 1$). The interaction of relatedness and verb form was not significant for either short or long IS, and also not significant when the analysis within IS length was split over SOAs.

To summarize, while the main effects of relatedness and length were still significant in the reanalysis with 10 items, the interaction of verb form, relatedness, and length was not obtained anymore. The number of 10 items was obviously too small to obtain the effect. On the basis of the present results, we cannot decide whether it would have shown up in a study with a low proportion of IS (as in Experiment 1), when the number of targets had been higher than 10.

Discussion

A phonological relatedness effect again occurred in this experiment. Moreover, the main effect of length was highly significant. While the interaction of SOA, verb form, relatedness, and length, which should occur according to the strict version of the predictions was not found, the interaction of relatedness, length, and verb form was significant on .05-level at SOA 0 ms. Long IS were more efficient than short IS in the past tense condition, where they matched the target's first syllable, while in the infinitive condition, short and long IS were equally efficient primes. This pattern had been predicted for the case that the segmental overlap effect and the syllable match effect co-occurred. For the past tense condition, segmental overlap effect and syllable match effect both predict a higher prime efficiency for long IS than for short ones, whereas for infinitives, the effects cancel each other: Short IS should be more efficient than long ones due to the syllable match effect, while long IS should be more efficient than short ones due to the segmental overlap effect. This result is promising. The fact that the different gain subjects had from short and long related IS interacted with the syllable structure of the target forms suggests that the syllable is a relevant unit in phonological encoding.

A problem is the point in time at which the interaction was obtained. The syllable match effect should be late in the encoding process. While the interaction was obtained at SOA 0 ms, the phonological priming effect at SOA 150 ms was still huge, but the interaction disappeared.

Experiment 3: Single Verbs, IS Cut Out of Target Form

The following experiment aimed to increase the chance to find the predicted effects. Two things were changed compared to the previous experiment. First, to get a better idea about the time range for phonological priming, that is, to find out which SOAs can be considered as late, four SOAs were tested instead of two. Secondly, the quality of the IS was changed. The underlying assumption for the previous experiments was that the spoken IS influence the encoding process at the segmental level. The IS had been spoken as syllables by a female speaker. However, it may well be the case that spoken IS have a direct influence at the level of syllable program nodes, or that a feature level is involved in encoding. The IS used so far perhaps had been too abstract, and IS that are as close as possible to the phonetic form of the targets might increase the effect. Therefore, the syllables were cut out of the target verb form in Experiment 3. This technique for creating experimental stimuli has been used already in the area of speech perception. Zwitserlood (1991) used it in a target string detection task. She took the first syllables of Dutch words like *buiging* (/bœʏ.ɣɪŋ/ "bowing") and *buigzaam* (/bœʏɣ.za:m/ "bowable") and found what can be interpreted as a syllable match effect in perception: /bœʏ/ was detected faster in /bœʏ.ɣɪŋ/ than in /bœʏɣ.za:m/, and /bœʏɣ/ was detected faster in /bœʏɣ.za:m/ than in /bœʏ.ɣɪŋ/. This effect disappeared when in a second condition the sequence /bœʏɣ/ was cut out of /bœʏ.ɣɪŋ/ and /bœʏ/ out of /bœʏɣ.za:m/. In the latter condition, /bœʏ/ was detected faster in /bœʏɣ.za:m/ and vice versa. These results indicate that the properties of the basis a string is taken of are relevant for the detection of that string when it is contained in longer words. Information about the syllabic structure of the base word and adjacent segments that do not belong to the target string can be extracted from the string by the speaker, probably by phonetic cues like coarticulation, pitch contours, and so forth. Syllables cut out of the target forms might thus be IS that are more likely to produce syllable match effects.

Method

Participants. In this experiment, 106 students took part, 27 at SOA -150 ms, 26 at SOA 0 ms, 25 at SOA 150 ms, and 28 at SOA 300 ms.

Stimuli. The same targets were investigated as in Experiment 2. The quality of the IS was different from the ones used in the previous experiment. Instead of being spoken as syllables, they were cut out from the whole target verb forms that had been produced by a female speaker. This was done with the speech processing software XWAVES. For instance, participants heard a /ko:/ cut out of the infinitive *koken* if they had to say the infinitive *koken*, and a /ko:/ cut out of the past tense form *kookte* if they had to say the past tense form *kookte*. The syllables were cut relying on the information provided by the visual time wave and simultaneous auditory control.

Design and procedure. Design and procedure were similar to Experiment 2, except that four SOAs were investigated (-150, 0, 150, and 300 ms).

Data analysis. Three subjects had to be excluded because of technical errors,

	Verb Form							
IS	Infinitive		Past Tense		Infinitive		Past Tense	
	SOA -150 ms				SOA 0 ms			
Short								
Related	694	(2.2)	696	(3.6)	731	(3.6)	724	(4.0)
Unrelated	819	(4.0)	824	(5.4)	882	(6.0)	891	(6.6)
Δ	125		128		151		167	
Long								
Related	705	(2.4)	703	(3.4)	727	(4.6)	754	(2.8)
Unrelated	848	(3.0)	841	(8.6)	901	(7.0)	914	(6.2)
Δ	143		138		174		160	
Pink Noise	772	(2.8)	761	(4.6)	810	(5.2)	817	(4.8)
	SOA 150 ms				SOA 300 ms			
Short								
Related	702	(1.8)	704	(2.2)	701	(2.6)	717	(3.2)
Unrelated	850	(4.8)	871	(8.2)	824	(6.4)	814	(7.2)
Δ	148		167		123		97	
Long								
Related	715	(1.8)	719	(2.0)	713	(1.8)	721	(2.4)
Unrelated	867	(6.8)	910	(6.6)	822	(8.6)	815	(7.8)
Δ	152		191		109		94	
Pink Noise	801	(2.8)	797	(4.8)	765	(5.6)	777	(6.2)

Table 6. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 3

Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses.

and three because they made many errors. Twenty-five students contributed to each SOA.

Results

The results are shown in Table 6 and Figure 7. The mean RT in the experiment was 786 ms. Infinitive verb forms yielded a mean RT of 783 ms, past tense forms of 789 ms. This difference was significant over items only ($F_2(1;19)=9.89$, $MSE = 868$, $p < .01$). Furthermore, a main effect of phonological relatedness was obtained ($F_1(1;96) = 430.68$, $MSE = 9313$, $p < .01$; $F_2(1;19) =$

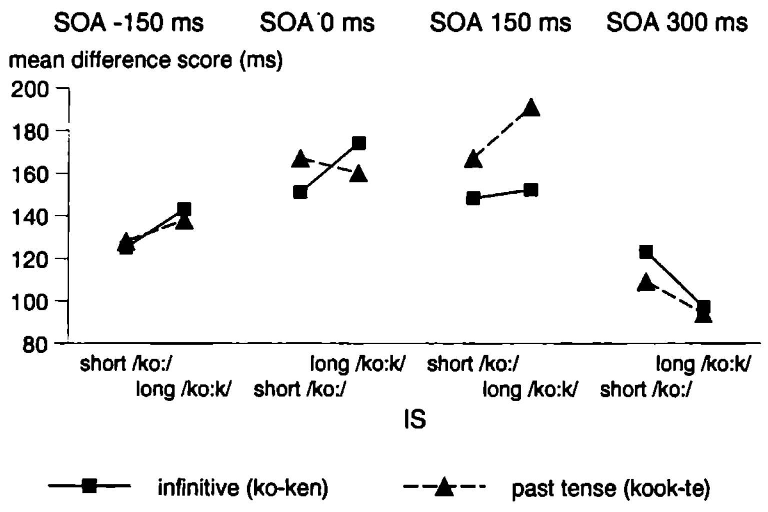


Figure 7. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Infinitive and Past Tense Forms at the Four SOAs in Experiment 3

303.23, $MSE = 10583$, $p < .01$). Related IS led to significantly faster reactions than unrelated IS (mean RT related = 714 ms, unrelated = 856 ms). The interaction of relatedness and SOA was significant ($F_1(3;96) = 4.2$, $MSE = 9313$, $p < .01$, $F_2(3;57) = 18.56$, $MSE = 1684$, $p < .01$). The priming effect was higher at SOAs 0 ms (163 ms) and 150 ms (165 ms) than at SOAs -150 ms (133 ms) and 300 ms (106 ms). Tests of simple effects confirmed that the effect of relatedness was significant for each level of SOA.

There was an effect for SOA in the item analysis only ($F_2(3;57)=38.34$, $MSE = 2354$, $p < .01$). Moreover, a significant main effect of IS length was observed

(F_1(1;96) = 44.5, *MSE* = 941, p < .01; F_2(1;19)= 14.97; *MSE* = 2255, *p* < .01). The interaction of length and SOA was significant on .05-level (F_1(3;96) = 2.79, *MSE* = 941, *p* < .05; F_2(3;57) = 3.53, *MSE* = 600, *p* < .02). While the effect of length was 16, 17 and 21 ms at SOAs -150, 0 and 150 ms respectively, it was only 4 ms at SOA 300 ms. Tests of simple effects showed that the main effect of length was significant at all SOAs, except of SOA 300 ms.

The interaction of the variables verb form, relatedness, and length was not significant (all *F* < 1). The interaction was analyzed for each SOA and was never significant. The interaction of relatedness and length was not significant, either. However, the interaction of relatedness and length was significant for the past tense forms at SOA 150 ms, but only on .05-level in the subject analysis (F_1(1;96) = 4.18, *MSE* = 852, p < .05; F_2(1;19)= 3.1; *MSE* = 919, *p* < .10).

Discussion

The consistent effect of priming for phonologically related IS over all SOAs shows that the time window investigated by these SOAs seems to be involved in phonological encoding in this experimental task. The amount of phonological priming increases from SOA -150 ms to SOA 0 ms, and drops substantially from SOA 150 ms to SOA 300 ms. At that late SOA, also the main effect of length disappeared. It was probably too late for the participants to perceive the whole IS. SOAs 0 ms and 150 ms showed the highest priming effects. A trend occurred at SOA 150 ms for long IS being better primes than short IS for past tense forms, while short and long IS primed infinitives equally well, as predicted for the case that syllable match and segmental overlap effect co-occurred. But the interaction of verb form, relatedness, and length, which had been obtained at SOA 0 in Experiment 2, did not show up in Experiment 3.

Experiments 2 and 3 differed in the type of IS used, and this may have caused the different results. While the IS in Experiment 2 were spoken as syllables (e.g., "koo") by a female speaker who had been instructed to produce the syllables as they occur in the target words (e.g., *koken*), the IS in Experiment 3 were cut out of the whole target word that had been spoken by the same speaker.

To find out whether the results of Experiments 2 and 3 differed significantly, Anovas were run over the combined data of Experiments 2 and the corresponding two SOAs of Experiment 3. The analyses included the variables 'verb form' (two levels: infinitive, past tense), 'phonological relatedness' (two levels: related, unrelated), 'length' (two levels: short, long), 'SOA' (two levels: 0 ms, 150 ms) and 'experiment' (two levels: Experiments 2 and 3). SOA and experiment were varied between subjects. The main effect of experiment was significant only in the item analysis (F_2(1;19) = 6.28, *MSE* = 4009, *p* < .03). Experiment 2 with spoken IS yielded a mean RT of 815 ms, while Experiment 3, which used IS cut out of the target forms, yielded slightly faster reactions (mean RT = 804 ms). The interaction of all variables was not significant over subjects and just failed significance in the analysis over items ($F_2(1;19) = 4.32$, $MSE = 1057$, $p < .06$).

The amount of phonological priming obtained in the two Experiments differed significantly ($F_1(1;96) = 175.11$, $MSE = 6459$, $p < .01$; $F_2(1;19) = 24.23$, $MSE = 1640$, $p < .01$). Related IS speeded up reactions more in Experiment 3, where they were cut out of the target form, than in Experiment 2, where they were spoken as syllables, both as compared to an unrelated baseline of corresponding length (priming effect Experiment 2 = 133 ms, Experiment 3 = 163 ms).

Furthermore, the Experiments yielded a main effect of length that differed significantly, but only in the subject analysis ($F_1(1;96) = 68.73$, $MSE = 2850$, $p < .01$): In both Experiments, participants reacted slower when they heard long than when they heard short IS. This difference was slightly larger in Experiment 3, where the IS were cut out of the target forms (19 ms) than in Experiment 2 that tested spoken syllable IS (15 ms).

Importantly, the interaction of 'experiment', 'verb form', 'relatedness', and 'length' was significant at SOA 0 ms in the items analysis ($F_2(1;19) = 6.84$, $MSE = 906$, $p < .02$). At this SOA, Experiment 2 with spoken syllable IS had yielded the pattern that was predicted if segmental overlap and syllable match effect co-occurred: Long IS had primed past tense forms more efficiently than short IS, while the prime efficiency of short and long IS the spoken syllable IS did not differ for infinitive targets.[3] Experiment 3, where the IS were cut out of the target form, lacked this interaction.

Since Experiments 2 and 3 were similar except for the type of IS that was used, the different results obtained in both experiments were probably caused by the different IS. The different methods by which the IS were created certainly affected their acoustic characteristics.

For instance, the spoken syllables in Experiment 2 were substantially longer than the IS that had been cut out of the verb forms in Experiment 3 (mean length spoken IS = 450 ms, IS cut out of target form = 310 ms). This is expected on the basis of phonetic research, which found the length of a particular syllable to decrease with an increasing number of syllables following it (Lehiste, 1970; Kohler, 1982).

[3]This interaction for the spoken syllable IS in Experiment 2 remained significant over items in the analysis over Experiments 2 and 3 ($F_2(1;19) = 4.87$, $MSE = 888$, $p < .04$).

A further difference between the experiments was that participants in Experiment 2 heard the identical IS (e.g., /ko:/) for past tense and infinitive targets, whereas participants of Experiment 3 heard different tokens of a stimulus in the two different verb form conditions, namely the one cut out of the verb form they currently produced. As a consequence, participants did not only hear two tokens of one stimulus, while participants in Experiment 2 heard only one, but the two versions of IS in Experiment 3 were in addition likely to differ systematically. Compare for instance, the short related stimulus /tra:/ cut out of the target *tranen* ("to water") and the target *traande* ("watered"). Due to Closed Syllable Vowel Shortening, which was mentioned in the Introduction as a possible phonetic cue to syllabic structure, a sequence /tra:/ spliced out of *traande* $(\text{tra:n})_\sigma(\text{də})_\sigma$ has a shorter vowel - because a consonant closes the syllable - compared to the vowel in the syllable /tra:/ from *tranen* $(\text{tra:})_\sigma(\text{nən})_\sigma$, where no consonant is in the syllable's coda. A further potential difference between the two versions of IS in Experiment 3 could be due to coarticulatory effects. If there is more coarticulation within than between syllables, as some phoneticians suggested (e.g., Fujimura & Lovins, 1978; see again the Introduction), the vowel in the sequence /tra:/ spliced out of the syllable /tra:n/ of *traande* will bear a high amount of nasality from the subsequent nasal caused by anticipatory coarticulation. In the sequence /tra:/ from *tranen*, the intervening syllable boundary could lower the degree of anticipatory assimilation to the nasal consonant. The mismatching syllable /tra:/ spliced from *traande* might thus nevertheless carry syllable match information because of coarticulation. The mismatching syllable might therefore resemble the matching syllable /tra:n/ such that not differences in priming are obtained. Experiment 2 lacks these differences, since the same abstract IS were used for both target verbs.

Furthermore, cutting the stimuli in Experiment 3 could in fact cause the exclusion of relevant syllabic information. This is not unlikely, given the insufficient knowledge about phonetic cues to syllabic structure. In spoken syllables, like those that had been used in Experiment 2, all information about syllabic status is by definition included in the stimulus. Therefore, the IS of the following experiments were all spoken as syllables.

Experiment 4: Encliticized Forms, 12.5% Related IS

The experiments reported so far tested single verbs as targets. A potential reason for the difficulty of obtaining the predicted effects might be that these forms are highly overlearned and processed too fast to pick up the effects. Moreover, a single lexical unit is identical with a prosodic word, and needs no process of prosodic word formation. In contrast to that, encliticized forms are unlikely to be highly overlearned, because clitics can attach to a variety of different words. Moreover they involve a process of prosodic word formation, since the syllable of the clitic function word has to be integrated into the preceding prosodic word frame.

Experiment 4 looked therefore at short strings of connected speech like *kook het* "cook it" $(\text{ko:})_\sigma(\text{kət})_\sigma$, where a schwa-initial weak function word obligatorily encliticizes to the verb. According to the model of phonological encoding, short IS (e.g., /ko:/), which correspond to the first syllable in encliticized forms like *kook het*, should be more efficient primes than long ones (e.g., /ko:k/) at a late SOA (= syllable match effect). At an earlier SOA, the model predicts the segmental overlap effect: Long IS should be more efficient than short ones, because they prime more segments. According to the less strict version of the predictions, where syllable match and segmental overlap effect co-occur, long and short IS should not differ in their prime efficiency, because long IS have more segmental overlap with the target, but short IS correspond to the target's first syllable.

Method

Pretest of the Stimuli

Participants. Twenty-nine students participated in the pretest.

Stimuli. There were certain phonological constraints on the verb to make sure that participants produced syllable structures like $(\text{ko:})_\sigma(\text{kət})_\sigma$: The verb had to end in a consonantal cluster: When a verb ending in a vowel, for instance *doe* ("do") /du:/, precedes a schwa-initial function word, this results in a hiatus. Dutch normally avoids a hiatus, for instance by inserting a homorganic glide. Thus, *doe het* ("do it") becomes $(\text{du:})_\sigma(\text{ʋət})_\sigma$. In this form, the verb stem still corresponds to the first syllable, which is not the case for $(\text{ko:})_\sigma(\text{kət})_\sigma$. We thus need a consonant, and even better a consonantal cluster. A cluster prohibits ambisyllabic consonants, which would occur after short vowels as e.g. in *tip het* "guess it" /tɪ(p)ət/, where /p/ is ambisyllabic. But not any cluster is fine: Coda clusters violating the sonority hierarchy were excluded, like /ks/ or /ts/, since the final coronal fricative is treated as an appendix by many phonologists, as has been discussed in the Introduction section of this thesis. Also, verbs ending in a voiced consonant underlyingly were excluded to avoid potential interactions with syllable-final devoicing. Furthermore, some clusters allow for optional schwa-insertion in coda position. The insertion is prohibited if the cluster is a nasal followed by a homorganic consonant (/mp/, /nt/, /ŋk/), or if the second consonant is a coronal obstruent (/s/ or /t/), or if the first consonant is no sonorant (like in /sp/, /sk/). In all other clusters, a schwa could be inserted, leading to an additional syllable, like for instance in *melk* ("milk"), which can be pronounced as [mɛlk] or as [mɛlək].

Because the materials were originally selected in order to conduct several follow-up-experiments with different personal pronouns, the verbs had to be combinable with *het* "it", *hem* "him", *ik* "I", and *ons* "us". The set of verbs that can go with all of these pronouns and in addition obey all phonological constraints was considerably small. As a result, some verbs had to be included that violated one of the constraints.

Two lists of words were created. The first list contained all monosyllabic verb stems of Dutch that could be combined with *ik, ons, hem*, and *het* ("I, us, him, it"). The second list contained the verbs of the first list that ended in a consonantal cluster. Verbs from the first list were used as fillers. From the second list, 24 verbs were chosen as candidates for experimental targets. The latter set (and in addition six filler pairs as practice items) was included in the pretest to control for the mean naming latencies and error rates for semantic-associates naming without interfering stimuli (see Appendix A3.1 for the materials and results).

Procedure. The students participated in single sessions. They were instructed to learn the list of nouns and semantically associated verbs by heart and to produce the verb as soon as the noun appeared on the screen. They were told to produce the verb's infinitive form first, and in the following blocks to combine the verb with the four function words. A participant started with six practice items, and then produced the 24 candidate verbs in the infinitive form. Then there was a sentence on the screen informing the participant which function word should follow the verb in the following block, e.g. *Zeg nu "hem" achter het werkwoord, zoals "test hem"* ("Now say 'him' after the verb, such as 'test him'"). The sentence appeared on the screen for 10 seconds, followed by a 1500 ms pause. Every block began with the six practice pairs, followed by the 24 verbs in a different randomization for every block and every participant. In total, each noun-verb pair occurred five times: Once in the citation form, and four times combined with the four different function words. There were four groups of participants who received the four function words in a different order.

Data analysis. Five participants had to be excluded because of a high error rate, and 24 good participants were analyzed, six for every participant group.

Results of the pretest. Each item occurred 120 times (five occurrences for each of the 24 participants). Six items had 10% errors or less, and 10 items had error percentages between 10% and 20%. In addition, to increase the number of items, four items were included that had more errors. Those four verbs were contained in one set, so that they could later be excluded from analysis easily, if they turned out to behave differently from the others.[4] In total, 20 targets were selected for the main experiment.

Main Experiment

Participants. Ninety-eight students performed in the experiment, 30 for SOA 0 ms, 34 for SOA 150 ms, and 34 for SOA 300 ms.

Stimuli. The targets consisted of a monosyllabic verb followed by the schwa-initial weak form of the pronoun *het* ("it") /ət/, for instance, *kook het* "cook it". Again, the IS were either phonologically related or unrelated. Both occurred in a long and a short variant. In contrast to the preceding experiment, no pink noise was included. The proportion of related IS was low.[5] 20 target verbs occurred in four IS conditions, resulting in 80 experimental trials, 40 of them with phonologically unrelated IS. Seventy fillers were included, so that 320 trials contained unrelated IS. The proportion of related IS was 12.5%. Fillers were combined with four unrelated IS. These were taken from two other filler stimuli, each of which contributed a short and a long IS. Both, fillers and experimental targets occurred thus four times.

The 20 experimental verbs were split into five sets of four verbs each. Always two of the verbs in a set mutually provided unrelated IS. In addition, every set had 14 fillers (see Appendix A3.2). No noun occurred as a trigger that could be combined with two verbs of a set equally well.

Design. The design comprised four types of IS (related short, related long, unrelated short, and unrelated long). Three SOAs were tested between subjects (0,

---

[4]For all of those items there was reason to believe that they would do better in the main experiment than in the pretest. The item *water - drenken* ("water - to water") was included because it was assumed that there would be less errors if it was not together in one group with *kan - schenken* ("pitcher - to pour"). The verb *wenken* "to wave" was included and the noun changed from *vinger* "finger" to *hand* ("hand"), which is an easier association, and therefore less errors were expected. The item *rechter - schorsen* ("judge - reverse") was close to the excluding criterion of 20% errors, had many voice clicks, but otherwise correct answers, and had the lowest RT of the remaining verbs. Finally, the item *vrees - duchten* ("fear - to fear") had the lowest RT among the remaining verbs. Because of the need for verb pairs matching in CV-structure, the item *hout - spalken* ("wood - to splint") had to be replaced by *ruzie - pesten* ("squabble - to torment").

[5]Chronologically, this experiment was run before Experiment 2, where the interaction had occurred with an increased proportion of 40% related IS.

150, and 300 ms).

Procedure. In the instruction, participants were told that after the practice block they should produce an imperative form with the verb by attaching the word *het* to it. They were told that they should produce the sequence as it occurs in natural speech. The experimentor demonstrated this by pronouncing three encliticized examples from the filler stimuli.

For every set, participants first had to produce the citation form of the verbs once, without auditory distractors. Then, the sentence *Zeg nu 'het' achter het werkwoord, zoals 'zoek het'* ("Now say 'it' after the verb, as in 'look for it'") appeared on the screen for 10 seconds, preceded and followed by a one-second pause. Then eight filler items gave participants the chance to get used to producing the encliticized forms.

The targets and fillers of the set followed, randomized in four blocks. Each block contained the four experimental targets and 14 fillers once, in different IS conditions. The first block contained the first experimental target with the related short IS, the second experimental target with the related long IS, the third experimental target with the unrelated short IS, and the fourth experimental target with the unrelated long IS. In addition, there were seven fillers with long and seven fillers with short IS. The second block contained the first experimental target with the unrelated long IS, the second experimental target with the related short IS, and so on. Again, the four blocks were rotated between participants.

Participants studied a sheet of paper with the nouns and the verbs in the infinitive form and were informed that the experiment consisted of a practice and an experimental block, which were interrupted by a short pause, in which more instructions would be given. Furthermore they were told that there were five lists of words to be learned in the experiment.

Five groups of participants differed with respect to the order in which the five sets were given. Group 1 started with set 1, group 2 with set 2 and so on. Every group consisted of five participants at each SOA.

The whole experiment took approximately one hour.

Data analysis. Twenty-five subjects were analyzed per SOA, surplus subjects were randomly excluded.

Results

Table 7 shows that set 5, which contained the items from the pretest that had not actually fulfilled the criteria for the main experiment, deviated from the other sets. The mean RT was substantially higher, as was the standard deviation, and participants made more errors than in the other sets. Therefore, set 5 was excluded from the analysis. This reduced the number of items from 20 to 16.

	set				
	1	2	3	4	5
RT	893	886	894	835	1001
StdDev	233	240	252	221	300
Errors	1.9	2.0	2.1	1.3	5.3

Table 7. Mean RTs, Standard Deviations, and Percentages of Errors for each Item Set in Experiment 4
Note. Values for RTs and standard deviations represent ms.

Table 8 and Figure 8 show the results with the 16 items.

The mean RT was 877 ms. The analysis revealed, first of all, a clear phonological priming effect ($F_1(1;72) = 117.42$, $MSE = 5092$, $p < .01$; $F_2(1;15) = 39.93$, $MSE = 9615$, $p < .01$): Related IS led to significantly faster response latencies than unrelated IS (mean RT related = 833 ms, unrelated = 922 ms). The interaction of SOA and relatedness was significant ($F_1(2;72) = 9.18$, $MS_e = 5092$, $p < .01$; $F_2(2;30) = 18.48$, $MSE = 1627$, $p < .01$). The priming effect was 93 ms at SOA 0 ms, 131 ms at SOA 150 ms, and 44 ms at SOA 300 ms, and was significant at all levels of SOA.

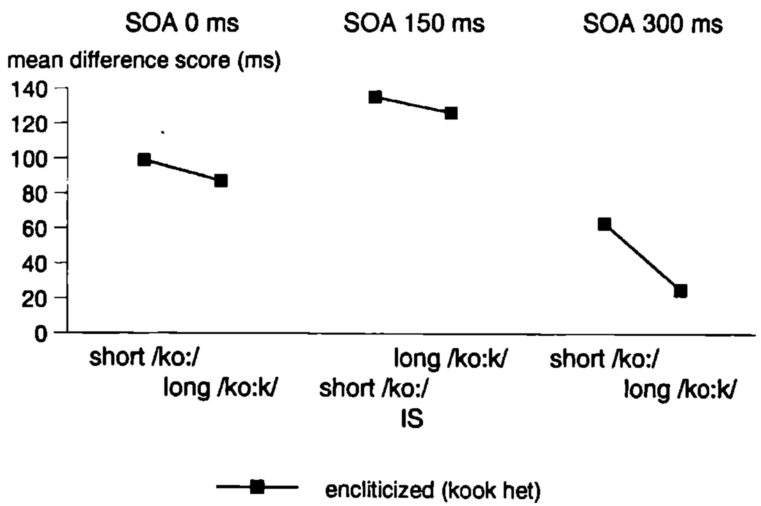


Figure 8. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Encliticized Forms at the Three SOAs in Experiment 4

IS	SOA 0 ms		150 ms		300 ms	
Short						
Related	819	(1.0)	789	(2.3)	852	(1.5)
Unrelated	918	(2.0)	924	(2.8)	915	(3.0)
Δ	99		135		63	
Long						
Related	840	(0.7)	807	(0.8)	889	(0.5)
Unrelated	927	(3.5)	933	(2.8)	914	(1.5)
Δ	87		126		25	

Table 8. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 4, Sets 1 to 4
Note. The values for RTs and difference scores represent ms, and percentages of errors appear in parentheses.

No other interaction was significant. The main effect of length gained significance over subjects only ($F_1(1;72) = 11.35$, $MSE = 1563$, $p < .01$). Long IS led to higher RTs than short ones (mean RT long = 885 ms, short = 870 ms). The effect of SOA was significant in the analysis over items only ($F_2(2;30) = 5.94$, $MSE = 2362$, $p < .01$). Although short IS were more efficient primes than long IS at each SOA, the interaction of relatedness and length was not significant ($F_1(1;72) = 3.09$, $MSE = 2313$, $p < .09$; $F_2(1;15) = 1.10$, $MSE = 4101$, $p < .31$). The interaction reached significance at no SOA, although at SOA 300 ms short IS surpassed long IS in prime efficiency by 38 ms (all $F < 1$, except at SOA 300 ms: $F_1(1;72) = 3.73$, $MSE = 2313$, $p < .06$; $F_2(1;15) = 2.74$, $MSE = 2017$, $p < .12$).

Discussion

In Experiment 4, encliticized forms replaced the single verb targets that had been tested in Experiments 1, 2, and 3. The clear effect of phonological relatedness that was obtained in Experiment 4 shows that phonologically related syllable IS efficiently prime not only single word responses elicited by a semantic-associate learning task, but also targets that contain more than one word. The main effect of length was significant, too, but only in the subject analysis.

Similarly to Experiment 1, the proportion of related IS was low in the present experiment, but it included a higher number of targets than the first experiment. No evidence was obtained for an early segmental overlap or a late syllable match effect. At SOA 300 ms, short IS were in fact more efficient than long IS, as predicted by the syllable match effect. But the difference did not reach significance.

However, for the case that segmental overlap effect and syllable match effect co-occur, they should cancel each other for targets with short first syllables: The former effect predicts long IS to be more efficient than short ones, while the latter effect predicts that short IS prime targets with short first syllables more efficiently than long ones.The prime efficiency of short and long IS should hence not differ for encliticized forms like *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ and it did indeed not differ in Experiment 4. Crucially, however, a control experiment is needed to investigate targets with a long first syllable. For these targets, long IS should be more efficient primes than short ones due to both segmental overlap and syllable match effect.

Experiments 5a,b: Encliticized Forms, 40% Related IS

Experiments 5a and 5b included encliticized targets with short and with long first syllables, respectively. In the course of the other experiments, only Experiment 2 revealed an effect that could be interpreted in line with the predicted segmental overlap and syllable match effects. This experiment included a high proportion of 40% phonologically related IS, which had been spoken as syllables by a female speaker. Experiments 5a and 5b ran under the same conditions as Experiment 2, but the targets were encliticized forms instead of single verbs.

The set of possible encliticized targets was larger in the present experiments than in Experiment 4, which had only included encliticized forms that could be combined with several pronouns (*hem* "him", *het* "it", *ik* "I", and *ons* "us"). In the present experiments, only the pronoun *het* ("it") was tested. Because of the larger set of possible targets, the experimental targets of Experiments 5a,b could be better controlled for the relation between cue noun and target, which was always an object-predicate relation. With these materials, participants responded substantially faster in the pretest of Experiment 5 than in the pretest of Experiment 4 (mean RTs = 798 ms and 935 ms, respectively; compare Appendix A1.1 and A4.1). Therefore, the SOA range was shifted in time compared to Experiment 4 by including SOAs -150 ms, 0 ms and 150 ms. SOA 150 ms was tested in Experiment 5a only.

Moreover, another type of encliticized form was introduced. Experiment 5a included targets similar to the targets in Experiment 4: An imperative verb form was followed by the pronoun *het*, e.g., *kook het* ("cook it"). In Experiment 5b, participants produced as a new type of encliticized forms the plural forms of imperative constructions, like *kookt het* ("cook (pl.) it"). This subtle change has consequences for the target's syllable structure: *kookt het* syllabifies $(\text{ko:k})_\sigma(\text{tət})_\sigma$. In contrast to the singular imperatives in Experiments 4 and 5a, where the first syllable was /ko:/ and corresponded to the short related IS, the first syllable of the targets in Experiment 5b was /ko:k/ and corresponded to the long related IS. Consequently, the predictions for the targets in Experiments 5a,b are similar to the predictions for infinitive and past tense forms in the single verb studies (Experiments 1 to 3). Singular imperatives should pattern like infinitives, and plural imperatives should pattern like past tense forms. For singular imperatives, segmental overlap effect and syllable match effect should affect the RTs for long and short IS, respectively, and therefore might cancel each other if they co-occurred. For plural imperatives, on the other hand, long IS should be more efficient than short ones regarding both, the segmental overlap and the syllable match effect. As a result the prime efficiency for long IS should be enhanced if the two effects co-occurred.

Method of Experiment 5a

Pretest of the Stimuli

Participants. Twenty-five students took part in single or double sessions.

Stimuli. The stimuli consisted of a list of 32 noun - verb pairs, where the noun and the verb had a predicate - object relation (for stimuli and results of the pretest see Appendix A4.1). Participants had to produce the target verb followed by the schwa-initial weak form of the pronoun *het* ("it") /ət/, for instance, *leer het* "learn it". In the first experiment with encliticized forms (= Experiment 4), the targets had been chosen from a set of Dutch monosyllabic verb stems that could be combined with four different function words. For the present experiment, it was clear that only the pronoun *het* "it (neuter)" would be tested. As a consequence, the set of possible monosyllabic target verbs was larger. To avoid schwa-epenthesis, they did not end in nonhomorganic coda clusters (except when the last C was a coronal). To exclude any possible effect of final devoicing, they did not end in voiced obstruents. To avoid ambisyllabic consonants, they all had long vowels or diphthongs or coda clusters. In all noun - verb pairs the noun was interpreted as the object of the verb. Due to gender agreement the noun was always a neuter noun. This led to more natural utterances than in the first experiment with encliticized forms.

Procedure. The 32 target verbs were divided into two sets, which participants learned by heart one after the other. One group of participants started with set 1, the other with set 2. The experimental items were produced three times, in three consecutive blocks within wich each items occurred once. There was no pause between the blocks. The randomization was different within every block and for every two participants. No two consecutive targets and nouns had the same onset or nucleus. A single trial looked like in the pretest of Experiment 4.

Data analysis. Two participants had to be excluded because of technical errors. Three participants were excluded who were the slowest within a group and made most errors, such that 10 participants were left in each group. The productions of the first block were excluded from the analysis.

Results of the pretest. The mean percentage of errors turned out to be low (1.9%), the mean RT was 796 ms. For the materials for the main experiment, three items with more than 10% errors were excluded, as were seven items exceeding RTs of 850 ms. One item had to be excluded because it ended in a nonhomorganic cluster that allows schwa-insertion. In two items, the noun was phonologically quite similar to the target verb. Those were also excluded. Of the remaining items, 20 were chosen that could be paired with respect to corresponding CV-structure. Of the 20 experimental verbs, four had a CVCC-structure (e. g., *pompen* "to pump"), four a CCVVC-structure (like *snoeren* "to bind"), and the remaining twelve were of the CVVC-type (for instance, *koken* "to cook"). In addition, six practice items were taken from the remaining verbs (see Appendix A4.2).

Main Experiment 5a

Participants. Ninety-two students participated in the experiment, 29 for SOA -150 ms, 35 for SOA 0 ms, 28 for SOA 150 ms.

Stimuli. The cue noun was phonologically unrelated to the verb in that they shared neither onset nor nucleus nor coda. IS occurred in five conditions (related short, related long, unrelated short, unrelated long, and pink noise). No fillers were used. The auditory IS were spoken as syllables by a speaker of Dutch.

Design. The design comprised five types of IS (related short, related long, unrelated short, unrelated long, pink noise), and three SOAs (-150, 0, 150 ms).

Procedure. The 20 target verbs were split into five groups of four verbs, that were randomized following the same procedure as in Experiment 2. Participants were tested in single or double sessions. A single trial had the same structure as in Experiment 4. The practice list consisted of two blocks within which each item (and the six practice items) occurred once. It was differently randomized for every two or three participants, obeying the same constraints as the experimental list. The sentence *Zeg nu 'het' achter het werkwoord, zoals "fluit 't"* preceded the experimental session.

Data analysis. Six participants were excluded because of technical errors, six participants were excluded that made 15% or more errors. Five surplus participants were randomly excluded, ending with 25 participants per SOA.

Results of Experiment 5a

Table 9 shows the results of Experiment 5, see also Figure 9. The mean RT in the experiment was 849 ms. The analysis revealed, first of all, a clear phonological priming effect ($F_1(1;72) = 206.06$, $MSE = 4646$, $p < .01$; $F_2(1;19) = 81.08$, $MSE = 9436$, $p < .01$). Related IS led to significantly faster reaction times than unrelated IS (mean RT related = 794 ms, unrelated = 907 ms). The interaction of SOA and relatedness was only significant in the analysis over items ($F_2(2;38) = 90.01$, $MSE = 851$, $p < .01$). The largest priming effect was obtained at SOA 0 ms

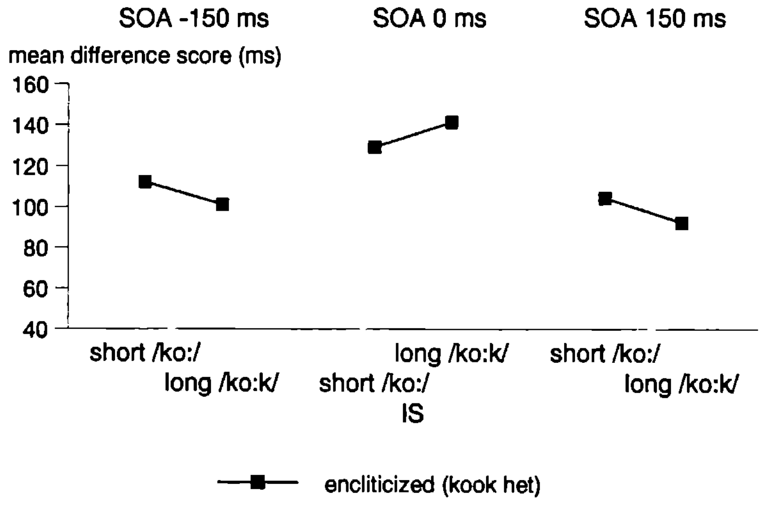


Figure 9. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Encliticized Forms at the Three SOAs in Experiment 5a

(135 ms), followed by SOA -150 ms (107 ms) and SOA 150 ms (97 ms). Tests of simple effects showed that the effect was significant at each level of SOA.

As in Experiment 2, 3, and 4 (in Experiment 4 only in the analysis over subjects), the main effect of length was significant ($F_1(1;72) = 18.24$, $MSE = 1363$, $p < .01$; $F_2(1;19) = 10.9$, $MSE = 1822$, $p < .01$). Long IS led to longer RTs that short ones (mean RT short = 842 ms, long = 860 ms). The interaction of SOA and length reached significance ($F_1(2;72) = 11.65$, $MSE = 1363$, $p < .01$; $F_2(2;38) = 10.48$, $MSE = 1216$, $p < .01$). Tests of simple effects revealed that the main effect of length was significant only at the first two SOAs, where RTs for short IS were faster than for long ones (36 ms difference at SOA -150 ms, 18 ms at SOA 0 ms), but not at SOA 150 ms, where the difference was only 11 ms. Possibly the IS came too late at the last SOA for the last segment to have an effect: The mean length of the IS was 510 ms, the mean RT in the experiment was about 850 ms. The same pattern had been found in Experiment 2, where the main effect of length had disappeared at SOA 300 ms.

	SOA					
IS	-150 ms		0 ms		150 ms	
Short						
Related	794	(4.6)	760	(5.2)	798	(4.2)
Unrelated	906	(7.4)	889	(8.0)	924	(8.2)
Δ	112		129		104	
Long						
Related	836	(3.8)	783	(4.0)	793	(4.2)
Unrelated	937	(5.4)	924	(8.2)	885	(7.8)
Δ	101		141		92	
Pink Noise	865	(6.6)	832	(4.2)	832	(7.8)

Table 9. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 5a
Note. The values for RTs and difference scores represent ms, and percentages of errors appear in parentheses.

The interaction of all factors (SOA, relatedness, and length) was not significant (all $F < 1$). The interaction of relatedness and length was not significant, either (all $F < 1$). Long and short IS were equally efficient primes at all SOAs, as tests of simple effects revealed.

Method of Experiment 5b

Experiment 5b was similar to the Experiment 5a, except that participants in Experient 5b were told to combine the verbs they had learned with the word *het*, as in *zoekt het* "searches it".[6]

Participants. Sixty-four students performed in the experiment, 34 for SOA -150 ms and 30 for SOA 0 ms.

	SOA			
IS	-150 ms		0 ms	
Short				
Related	820	(6.0)	770	(8.0)
Unrelated	925	(12.3)	887	(9.7)
Δ	105		117	
Long				
Related	822	(5.7)	771	(4.7)
Unrelated	894	(9.7)	891	(8.3)
Δ	72		120	
Pink Noise	848	(7.0)	813	(8.0)

Table 8. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 5b
Note. The values for RTs and difference scores represent ms, and percentages of errors appear in parentheses.

Data analysis. Five participants had to be excluded due to technical errors, seven because they made more than 15% errors, and one surplus participant was excluded. The experiment was run with the same materials to keep the degree of difficulty similar to Experiment 5a. However, 5 of the 20 target verbs had a stem ending in a /t/, like e.g., *fluiten* "to whistle". In contrast to stems ending in other sounds, those forms do not get a second /t/ due to third person singular inflection, see *kook+t het*, but *fluit het.* As a consequence, the imperative and third person singular forms do not differ. The first syllable of *fluit het* is /fləʏ/ and corresponds

[6] Furthermore, the noun of one semantic-associate filler pair was changed (*racisme*became *onweer* to cue *haten*). The reason was that after having performed the experimental list of this experiment, participants had to perform a block with different IS conditions for exactly the same list (see Experiment 7), only that the verb *haten* was an experimental target in that second experiment and therefore needed a noun that was not phonologically related.

to the short, and not the long related IS.

Therefore, the five targets ending in /t/ were excluded from the analyses, and also the targets with which they mutually provided the unrelated IS of corresponding length. The analyses were run over the remaining 12 items.

Results of Experiment 5b

The mean RT in the experiment was 844 ms, this is 5 ms faster than in the preceding experiment. The different form that participants had to produce thus did not cause a change in mean RTs. Table 10 shows the results, see also Figure 10.

The effect of phonological relatedness was significant ($F_1(1;48) = 134.62$, $MSE = 3987$, $p < .01$; $F_2(1;11) = 85.39$, $MSE = 3013$, $p < .01$), but not the main

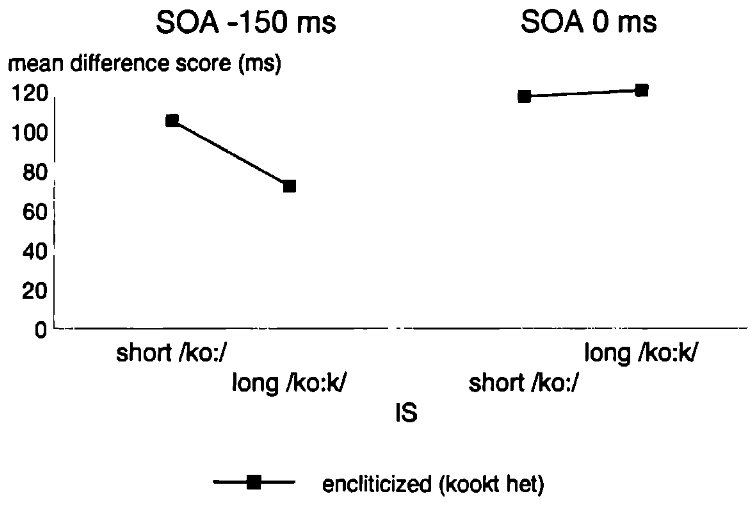


Figure 10. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Encliticized Forms in Experiment 5b

effect of length. SOA was again significant over items ($F_2(1;11) = 26.60$, $MSE = 1148$, $p < .01$). Short IS were more efficient than long ones at an early SOA, while long IS were more efficient than short ones at a late SOA. While the former pattern was unexpected, since short IS should be less efficient than long IS because they show less segmental overlap and do not match with the target's long first syllable, the findings for the late SOA were predicted and could be either caused by a syllable match effect or by the co-occurrence of segmental overlap and syllable match effect. However, the interaction of relatedness and length was not significant, nor was it significant in tests of simple effects at each SOA.

Discussion

The present experiments were designed parallel to Experiment 2, where an effect had been found that could be interpreted as a syllable match effect. The IS were spoken as syllables, and the percentage of related IS was high (40%). Unlike Experiment 2, the experiment tested encliticized forms, for which the syllable structure is not likely to be overlearned, as it might be the case for the single verbs, and has to be created during speaking. These changes should increase the chance of finding the predicted effects.

Experiment 5a tested encliticized forms of the type *kook het* ("cook (sg.) it"). A a clear phonological priming effect and a main effect of length indicated that the experiment was sensitive to phonological processing and to the different length of the IS, while the prime efficiency of long and short IS did not differ at the different SOAs. This result is in harmony with the less strict version of the predictions, according to which segmental overlap and syllable match effect co-occur: The syllable match effect should show short IS, which correspond to the first syllable of encliticized forms, to be more efficient than long IS, while long IS should be more efficient than the short IS, due to the larger segmental overlap.

Experiment 5b tested plural forms of encliticized forms, like *kookt het* ("cook (pl.) it"). Although long IS were more efficient than short ones at SOA 0 ms, which is in harmony with either the syllable match effect or the co-occurring segmental overlap and syllable match effect, no effect or interaction turned out to be significant, apart from a clear phonological priming effect. A possible problem with this experiment is the fact that the target sequence *kookt het*, though grammatical in Dutch, is slightly marked pragmatically. It can either be interpreted as the plural imperative ("cook it" spoken to more than one addressee) or as the imperative in the polite form. Both forms are homophonous. Participants reported that in the former case, they would rather use the singular form (*kook het*), also when addressing more than one person. And for the latter interpretation, that they would never say a command like this to somebody they have to address with the polite form. This difficulty is also reflected in the mean error rate, which was higher in the present study (7.9%) than in the preceding experiment (5.8%).

Experiment 6: Morphology Control

In Experiment 2, long IS were more efficient than short IS in priming past tense forms at SOA 0 ms, while prime efficiency of short and long IS did not differ for infinitives. This pattern had been predicted for the case that segmental overlap effect and syllable match effect co-occurred.

One possibility is that the obtained result was not due to syllabic structure at all, but caused by another variable. Two variables are confounded with IS length: Firstly, long IS always corresponded to the verb's stem morpheme, while short IS did not, compare, for instance, /ko:/ and /ko:k/ in $kook_V$+en_{affix}, $kook_V$+te_{affix}, and $kook_V$ het_{clitic}. A possible syllable effect might thus have been confounded with a morphological effect. Secondly, long IS were always lexical units, while the short IS mostly were not. This might have evoked a general lexical effect, for instance, a lexical inhibition effect. To test whether the effects obtained so far were due to morphological or lexical rather than phonological processing, two control experiments were run.

The first control experiment tested morphological processing. Presupposing that morphological and phonological effects are additive, and under the assumption that a morphological effect is facilitative, the morphological effect should work in the same direction as the segmental overlap effect, since the long IS corresponded to the stem morphemes. The syllable match effect and the morphological effect, on the other hand, should work against each other in the infinitive verb form condition (e.g., *koken*: long /ko:k/ = no syllable match, but morphological match, short /ko:/ = syllable match, but no morphological match), while in the past tense verb forms, the long IS should gain from both, morphological and phonological priming (e.g., *kookte*: long /ko:k/ = syllable match and morphological match). This provides an alternative explanation for the pattern at SOA 0 ms in Experiment 2, where the interaction of relatedness and length was only significant in the past tense forms, but not in the infinitives at SOA 0 ms. This asymmetry had been explained by assuming that segmental overlap effects and syllable match effect co-occurred. It might as well be caused by the co-occurrence of a syllable match and a morphological priming effect. Furthermore, if the morphological effect was much more elaborated than the syllable match effects, it could have hidden the latter, since long IS would be more efficient for both target verb forms.

An inhibitory morphological effect, on the other hand, should cancel the segmental overlap effect for both verb forms, or it should cancel the syllable match effect for the long IS in the past tense condition (e.g., *kookte*: long /ko:k/ = syllable facilitation and morphological inhibition). For infinitives, it should enhance the difference between long and short IS, since long IS get morphological inhibition and short IS get syllable match facilitation.

Little is known about morphological processing in language production. Drews and Zwitserlood (in press) used contiguous orthographic priming in a naming task. Whether the target was immediately preceded by a related prime or by a prime that was in addition also morphologically related did not have an effect: Both speeded up naming latencies compared to a neutral and an unrelated baseline condition. The authors accounted for this effect by assuming a direct pronunciation route that does not involve lexical processing. This grapheme-phoneme-correspondence account cannot be adopted for auditory priming experiments.

Hypothesis. The question under investigation is whether the auditory presentation of IS that correspond to a stem morpheme of a morphologically complex word influences the production of that complex word opposed to a target that has no morphological relation with the IS. For instance, does the morpheme *boot* ("boat") /bo:t/ given as an interfering stimulus influence the response latencies for the morphologically complex word *boot+en* "boats" $(\text{bo:})_\sigma(\text{tən})_\sigma$ in a different way than for the simple word *boter* "butter" $(\text{bo:})_\sigma(\text{tər})_\sigma$, when compared to the unrelated stimulus *paar* ("pair") /pa:r/ ? Notice that /bo:t/ has the same amount of phonological overlap for both complex and simple words and should be of same efficiency as far as phonology is concerned.

If /bo:t/ shows the same prime efficiency for both complex and simple targets compared to the unrelated /pa:r/, we can conclude that phonological, not morphological relatedness is relevant. If on the other hand the prime efficiency of /bo:t/ differs for complex and simple targets, the IS influence morphological units.

To make sure that the difference in prime efficiency is not a mere target complexity effect, short IS were included that corresponded to the targets' first syllable, like /bo:/. The related short IS should be more efficient than the unrelated short IS for both, complex and simple words alike. The same task was used as in the preceding experiments.

Method

Pretest of the Stimuli

Participants. Twenty-five students participated in the pretest. Some of them had participated in one of the preceding experiments.

Stimuli. In contrast to the preceding experiments, the targets were not verb forms, but noun pairs like *boter* "butter" and *boten* "boats". The targets had initial segmental overlap and a similar syllable structure, but different morphological structure. One target was always morphologically simple, and the other was a morphologically complex plural form. Twenty-three complex - simple noun pairs with overlapping first syllable plus onset of second syllable were investigated in a pretest, resulting in 46 semantically associated cue noun - target noun pairs.

Procedure. The targets were distributed over two sets. The first set contained 11 simple and 12 complex words, and the second 12 simple and 11 complex words (see also Appendix A5.1). The two members of a simple - complex pair occurred in two different sets. In the pretest, participants named the singular and the plural forms of both complex and simple targets. In the plural condition, the cue on the screen appeared in the plural, and in the singular condition, it appeared in the singular form. For some words, there was no plural, like e.g. for *boter*. Those were excluded in the plural condition. The experiment consisted of four parts, which were performed in different order by four participant groups: One group of participants started with the singular of set 1, followed by the plural of set 1, the singular of set 2, and the plural of set 2. One group started with the plural of set 1, singular of set 1, plural of set 2, singular of set 2. A third group started with the singular of set 2, followed by the plural of set 2, singular of set 1, plural of set 1, and so forth. Between parts were pauses, where participants received a new sheet of paper with the pairs in the appropriate number condition, which they learned. Each part started with three practice trials. Within each part, every target was presented three times in three concesutive blocks. A single trial looked like in the previous pretests. Students participated in sessions by one or two. The duration of the pretest was about 30 minutes.

Results of the pretest. The first occurrence of every target was excluded from the analyses. The percentage of errors was 6.7%. Five participants who had a higher percentage of errors than 6.7% were excluded, ending with 20 participants. For the singular simple words (the *boter*-words) and the plural complex words (the *boten*-words), the number of errors and the mean RTs for every target word were calculated (see Appendix A5.2). One pair had to be excluded because of a high number of errors, two other words because of high RT values. Furthermore, six pairs had to be excluded because the difference in RTs for the simple and the complex member of a pair exceeded 100 ms.

Fourteen simple-complex target pairs remained as experimental targets for the main experiment. Their mean lemma frequencies were 6.42 per million for the simple words and 7.41 per million for the complex words, the mean RT values were 811 ms for the simple, and 810 ms for the complex words.

The six pairs for which the RTs for the simple and the complex member differed by more than 100 ms were nevertheless included as fillers. This was done because the experiment was to be a control for Experiment 2, which had been run with 20 targets. The 14 targets plus 6 fillers provide an equally sized set of materials to be learned by the participants for one experimental set.

Main experiment

Participants. Ninety-five students took part in the experiment, 31 each for SOA 0 ms and 150 ms, and 33 for SOA 300 ms. Some of them had also participated in one of the preceding experiments.

Stimuli. The target words were morphologically simple or complex nouns with initial phonological overlap, like *boten* (boats) - *boter* (butter). These were primed by the five types of IS. Importantly, long related IS (e.g., /bo:t/) corresponded to the stem of complex targets (e.g., *boten*), but not of simple targets (e.g., *boter*). Short IS (e.g., /bo:/), on the other hand, were the first syllable of both words, but not a morphological unit. Table 11 shows an example for targets and IS.

Targets:			
Morphologically Complex:	*boten*	$(bo:)_\sigma(t\partial n)_\sigma$	"boats"
Morphologically Simple:	*boter*	$(bo:)_\sigma(t\partial r)_\sigma$	"butter"
Interfering Stimuli:			
	short	long	
Phonologically Related:	bo:	bo:t	
Phonologically Unrelated:	pa:	pa:r	
Pink Noise			

Table 11. Example for Targets and IS Conditions in the Morphology Control Experiment

Procedure. Contrary to the pretest, the variation of target number was excluded in the main experiment. The complex targets were produced in the plural form (e.g., *boten*), the simple targets in the singular form (e.g., *boter*). The cue word always appeared in the singular form to avoid possible priming by the plural inflection. The singular case has no overt marker in Dutch.

Targets and fillers were distributed over two sets of 20 semantically associated pairs each. One set contained the 20 simple and the other the 20 complex words. In addition, every set contained four practice items (see Appendix A5.3). Half of the participants started with the complex words and after a short pause continued with the simple ones. After another pause, they again had to name the simple words, followed by the complex ones. The other half of the participants performed the sets in the opposite order.

Before an experimental list was presented, participants completed a separate practice list, which consisted of two differently randomized blocks, whithin which every target word occurred once. At the respective SOA, pink noise was presented in the practice list. Every presentation of an experimental list started with six practice trials, in which all IS conditions occurred once (pink noise twice). The 20 words were randomized in five blocks, following the same procedure that had been used in the preceding experiments. Not more than five participants received the same randomization. The experiment lasted about 50 minutes.

Data analysis. Ten participants were excluded because they made more than 10% errors. Of the remaining 85 participants, ten were randomly excluded, until there were 25 participants left per SOA.

An Anova over subjects and two Anovas over items were run over the mean reaction times of the 14 experimental targets: The Anovas comprised the variables 'SOA' (three levels: 0, 150 and 300 ms), 'morphology' (two levels: complex (*boten*) and simple (*boter*)), 'relatedness' (two levels: related and unrelated), length (short and long), and 'repetition' (two levels: first and second exposure). One Anova over items treated morphology as a within-items factor; the second as a between-items factor.[7] SOA was tested between subjects.

Results

The mean RT for simple targets was 805 ms, for complex targets 794 ms. The effect of morphology was significant on .05-level over subjects, but not significant over items ($F_1(1;72) = 4.7$, $MSE = 6946$, $p < .04$).

Table 12 and Figure 11 show the results of the Morphology Control Experiment. The analyses revealed a significant effect of relatedness ($F_1(1;72) = 227.35$, $MSE = 10681$, $p < .01$; $F_2(1;13) = 97.73$, $MSE = 13917$, $p < .01$). The effect of SOA gained significance over items ($F_2(2;26) = 7.47$, $MSE = 4633$, $p < .01$).

[7]The morphologically complex - simple pairs, like *boten* - *boter*, involved different words, what suggests a between-items analysis for the factor morphology. However, the words were not independent, but matched according to the phonology and also with respect to the results in the pretest, where pairs with very different naming latencies had been excluded. This suggests a within-items analysis of the factor morphology. The F_2-values in the text come from the Anova with morphology as a within-items factor. The values from the analysis over items with morphology as a between-items factor are reported in Appendix B.

There was one significant interaction, which involved the variables relatedness and length ($F_1(1;72) = 15.79$, $MSE = 1721$, $p < .01$; $F_2(1;13) = 5.41$, $MSE = 2794$, $p < .01$). While related short IS led to slightly slower RTs than related long ones (mean RT related short = 759 ms, long = 752 ms), the opposite was the case in the unrelated IS (mean RT unrelated short = 840 ms, long = 852 ms). This

	Morphology							
IS	Complex		Simple		Complex		Simple	
	SOA 0 ms				SOA 150 ms			
Short								
Related	755	(1.6)	779	(2.4)	752	(1.1)	764	(3.7)
Unrelated	846	(3.3)	851	(2.9)	849	(4.0)	845	(4.9)
Δ	91		72		97		81	
Long								
Related	754	(1.7)	776	(2.6)	735	(1.9)	760	(2.7)
Unrelated	859	(5.4)	871	(5.3)	857	(4.1)	869	(5.7)
Δ	105		95		122		109	
Pink Noise	792	(1.9)	816	(3.3)	786	(2.3)	811	(3.6)
	SOA 300 ms							
Short								
Related	747	(2.3)	758	(3.0)				
Unrelated	823	(5.0)	824	(6.1)				
Δ	76		66					
Long								
Related	743	(1.9)	746	(3.3)				
Unrelated	826	(4.7)	829	(5.6)				
Δ	83		83					
Pink Noise	786	(3.6)	773	(3.4)				

Table 11. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors in Experiment 6
Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses.

results in a slightly higher prime efficiency for long IS (100 ms) compared to short ones (91 ms). Tests of simple effects showed that the interaction of relatedness and length was only significant on .05-level over subjects at SOA 0 ($F_1(1;72) = 4.64$, $MSE = 2059$, $p < .04$), and was significant at SOA 150 ($F_1(1;72) = 10.31$, $MSE = 1721$, $p < .01$; $F_2(1;13) = 10.32$, $MSE = 965$, $p < .01$). At that SOA, the priming effect for short IS, which correspond to the targets' first syllable, was 27 ms smaller than for long IS. This is the opposite pattern than predicted by the syllable match effect. Tests of simple effects within the simple and complex targets at each SOA made sure that the interaction of relatedness and length never reached significance.

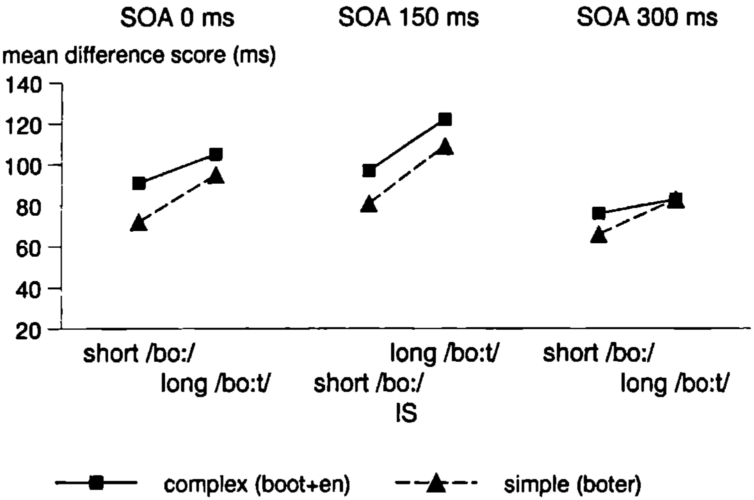


Figure 11. Mean Difference Scores (Unrelated - Related) for Short and Long IS Priming Morphologically Complex and Simple Targets in Experiment 6

At all SOAs, complex targets could be primed slightly better than simple targets, but this was the case for short as well as for long IS. This tendency did not gain significance: The interaction of relatedness with morphology was tested for each level of length at each SOA and was never significant, showing that the prime efficiency of long primes did not differ for simple or complex targets at each SOA, and the same held for short primes. Although long IS were morphologically related to the complex targets (e.g., /bo:t/ "boat" for target *boot+en* "boats"), while short IS were not, this obviously did not affect prime efficiency. The interaction of relatedness, length, and morphology was not significant at any SOA.[8]

	Morphology	
	complex	simple
Short		
Related	752	767
Unrelated	840	840
Δ	88	73
Long		
Related	744	761
Unrelated	847	856
Δ	103	95

Table 13. Mean RTs and Mean Difference Scores (Unrelated - Related) Within the Morphology Conditions in Experiment 6
Note. The values for RTs and difference scores (Δ) represent ms.

A test of simple effects on the interaction of relatedness and length at the two levels of morphology (collapsed over the three SOAs) showed a significant effect for the simple words like *boter* ($F_1(1;72) = 9.28$, $MSE = 2059$, $p < .01$; $F_2(1;13) = 4.78$, $MSE = 2239$, $p < .05$). Table 13 shows this result. Long IS were more efficient than short IS in the simple words, where the priming effects were 95 ms for long IS and 73 ms for short IS. For complex words like *boten*, the effect was in the same direction (88 ms priming for short IS, 103 ms for long IS), but was only significant over subjects on .05-level ($F_1(1;72) = 4.18$, $MSE = 2158$, $p < .05$; $_2(1;13) = 2.79$, $MSE = 1777$, $p < .12$). However, one has to keep in mind that the three-way interaction of morphology, relatedness, and length had not been significant.

Discussion

Experiment 6 investigated the influence of morphology by systematically varying the morphological and phonological relatedness of IS and target words. The presentation of phonologically related IS yielded a highly significant facilitative

[8]All F-values < 1, except at SOA 0 over items (morphology treated as within-items factor): F_2 (1;26) = 2.66, $MSE = 1712$, $p < .1$.

effect at all SOAs (0, 150, and 300 ms), and for both morphologically simple and complex targets. Moreover, the factors IS length and relatedness interacted significantly: Long IS were better primes than short ones. This also shows that participants perceived the difference between short and long IS, although the main effect of length did not reach significance. The interaction of SOA, relatedness, and length, was not significant (all $F < 1$). Importantly, with respect to the morphology issue, neither was the interaction of SOA, morphology, relatedness, and length significant, nor the interaction of relatedness, length, and morphology at either SOA. Long and short IS primed complex and simple targets equally efficiently. IS that were only phonologically related to the target produced the same facilitative effect than IS that were both phonologically and morphologically related.

The aim of this experiment had been to make sure that effects obtained in previous experiments were not due to morphological match of the IS. It can be concluded that, at least for the SOAs tested, morphologically related IS behaved like pure phonological primes in the Morphological Control Experiment.

Experiment 7: Lexical Control

A second control experiment investigated potential lexical effects that could have influenced the results of the preceding experiments. Lexical effects could evolve because long IS were words in most cases, while short IS were nonwords most of the times. This could lead to a lexical interference effect: A node in the lexical network is activated when a word is presented, which is the case for long, but not for short IS. This activated node could inhibit the target word's activation and slow down RTs for the long IS conditions. The lexical effect would, for instance, cancel out the segmental overlap effect, if both effects were additive and equally efficient.

In the Lexical Control Experiment, participants produced infinitive verb forms. Instead of short and long IS, they heard IS that were either words or pseudowords. Word and pseudoword IS were of equal length and only differed in their last segment. For example, the related word stimulus for the target verb *koken* ("to cook") was *koor* ("choir"), while the related pseudoword stimulus was *koom*.

Method

The experiment was run following Experiment 5b. The same semantically associated noun - verb pairs were used.

Participants. The same students participated as in Experiment 5b.

Stimuli. Participants produced the infinitive form of the verbs, for instance, *koken* "to cook". At SOAs 0 ms and 150 ms, which had been the SOAs tested in Experiment 2, syllable IS were presented that were again phonologically related or unrelated and spoken by the same speaker as in Experiment 5b. Within both relatedness conditions, IS did not differ in length, but in their lexical status. They were either a word or a pseudoword. Both related IS, the word and the pseudoword, shared the segments of the target's first syllable and differed in one additional segment, for instance, *koor* "choir" and *koom* for the target verb *koken* "cook". As in the previous experiments, pairs of targets that were matched in CV-structure mutually provided the unrelated IS, for example, *vies* "dirty" and *vien* (pseudoword) from *vieren* ("celebrate") for the target *koken.* A similar number of each word and pseudoword IS ended in stops (five words, six pseudowords), fricatives (both five), nasals (both four) and liquids (four words, three pseudowords) to make sure that the IS did not differ in perceptual saliency. Word IS had a mean length of 563 ms, pseudoword IS of 559 ms. In addition, there was a pink noise corresponding in length to the mean length of the IS in the experiment (see Appendix A6.1).

For 4 of the 20 target verbs of Experiment 5b, no phonologically related pseudowords and words could be found. Two targets were added that had been fillers in Experiment 5b, resulting in 18 targets for the Lexical Control Experiment.

Design. The design comprised two types of relatedness (related and unrelated), and two types of lexical status (word and pseudoword), which were crossed, and two SOAs (0 and 150 ms), which were tested between subjects.

Procedure. There was a short (5 s) pause between Experiment 5b and the Lexical Control Experiment. The experiment started with the sentence *Zeg nu het hele werkwoord, zoals "tinten"* ("Now say the whole verb, such as 'tinten'"), followed by the attention sign and 13 practice trials to give participants the chance to get used to the new type of IS and the new SOA. For participants who had performed in SOA -150 ms in experiment 5b, the IS were presented at SOA 0 ms in the Lexical Control Experiment, and for participants of SOA 0 ms in Experiment 5b, the IS were presented at SOA 150 ms, such that for both participants groups the SOA changed between the two consecutive experiments. The practice trials contained the two items that had previously been fillers each twice (combined with pink noise as an IS), and three additional fillers occurred such that participants heard all IS conditions at least once.

Targets were again randomized in five consecutive blocks in which each item occurred once, following the same procedure as in the preceding experiments. The 18 targets were divided into five item groups, three of which contained four and two contained three items. The items of an item group shared the IS condition within a block, and each item occurred once in every IS condition during the experiment.

Five groups of five participants each started with different blocks. Within each participant group, two or three participants performed the same randomized list.

Data analysis. Two participants were excluded due to technical errors, six because of a high number of errors, and six surplus participants, such that 25 participants were left per SOA.

Results and Discussion

As shown in Table 14, phonologically related IS again facilitated responses (mean RT related = 789 ms, unrelated = 852 ms). The phonological priming effect was highly significant ($F_1(1;48) = 61.28$, $MSE = 3166$, $p < .01$; $F_2(1;17) = 37.01$, $MSE = 3779$, $p < .01$). The interaction of SOA and relatedness was not significant. Most importantly, the lexical status of the IS had no effect (all $F < 1$), nor did it interact with any other variable. Words and pseudowords provided an equal amount of priming. The mean RT values differed in 3 ms (words = 822 ms, pseudowords = 819 ms), and within the different relatedness conditions and SOAs, the difference

	Lexical Status of IS			
IS	Word		Pseudoword	
	SOA 0 ms			
Related	808	(6.0)	801	(3.3)
Unrelated	864	(6.4)	863	(6.4)
Δ	56		62	
	SOA 150 ms			
Related	772	(2.9)	777	(5.3)
Unrelated	846	(5.6)	836	(5.8)
Δ	74		59	

Table 14. Mean RTs, Mean Difference Scores (Unrelated - Related), and Percentages of Errors for Word and Pseudoword IS in Experiment 7

Note. The values for RTs and difference scores (Δ) represent ms, and percentages of errors appear in parentheses. Mean RTs for pink noise IS (with error percentages in parentheses) were 841 ms (4.9) at SOA 0 ms and 795 ms (4.0) at SOA 150 ms.

never exceeded 10 ms.

To conclude, the Lexical Control Experiment showed that whether IS are lexical units or not has no effect on the response latencies. This issue arose because in the preceding experiments, the length of the IS was confounded with a lexical variable: Long IS were always lexical units, while short IS were not most of the times. The results of the Lexical Control Experiment make it unlikely that the differences in RTs that had been obtained for long and short IS in the preceding experiments were due to their different lexical status.

In sum, the two control experiments ruled out the possibility that the prime efficiency of short and long IS depends on morphological or lexical variables. If short and long IS differ in prime efficiency, the locus of the difference presumably is the phonological level.

Reanalysis with Pink Noise as a Baseline

In all experiments reported so far, the RTs obtained with related IS were subtracted from the RTs that had been obtained with unrelated IS of corresponding length. In this section I will discuss alternatives to the unrelated baseline. The RTs obtained in the related IS conditions can be either compared to the RTs that were obtained when participants heard a pink noise stimulus, or they can be analyzed without any baseline condition.

As a first alternative, one might look at the gain that related IS yielded when compared to the pink noise IS, instead to compare the response latencies for related IS with the latencies obtained for unrelated IS of corresponding length. In all experiments, related IS led to faster RTs than pink noise. Moreover, related short IS often yielded more facilitation as compared to the pink noise baseline than

	SOA		
IS	-150 ms	0 ms	150 ms
Δ PinkNoise - Related			
Short	71	71	34
Long	29	49	39

Table 15. Mean Mean Difference Scores (Related - Pink Noise) in Experiment 5a

Note. The values for difference scores (Δ) represent ms.

related long IS. This difference was most pronounced in Experiment 5a, where the targets were encliticized forms of the type *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$, and the short IS corresponded to the targets' first syllable. The effects that result from a comparison of the RTs for short IS with the RTs for the pink noise condition are shown in Table 15 (see also Figure 13).

At SOAs -150 ms and 0 ms, related short IS yielded more facilitation than long ones (42 ms and 23 ms, respectively). Anovas were run over the difference scores of the related IS conditions from pink noise. They showed a significant effect of length ($F_1(1;72) = 13.85$, $MSE = 1105$, $p < .01$; $F_2(1;19) = 7.25$, $MSE = 1693$, p

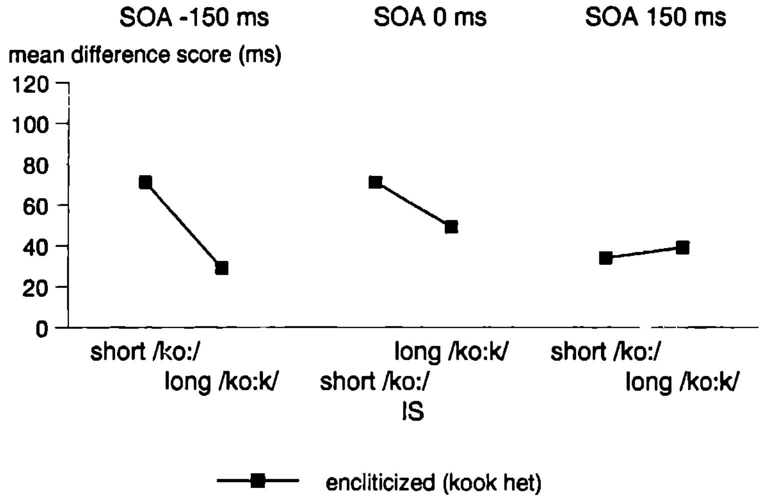


Figure 13. Mean Difference Scores for Short and Long Related IS from Pink Noise in Experiment 5a

$< .02$), and an interaction of SOA and length ($F_1(2;72) = 6.07$, $MSE = 1105$, $p < .01$; $F_2(2;38) = 6.54$, $MSE = 828$, $p < .01$). Tests of simple effects showed that the difference in facilitation for short and long related IS was significant at SOA -150 ms and SOA 0 ms, but not at SOA 150 ms. A posteriori tests (Dunnett's tests) revealed that the RTs in the related conditions differed from the pink noise baseline significantly in almost all conditions (all $p < .01$, except SOA -150 ms, related long condition, which was significant on .05-level over items and on .01-level over subjects).[9] At SOAs -150 ms and 0 ms, short related IS, which corresponded to the target's first syllable, provided more facilitation than long IS. This result could be interpreted as a syllable match effect.

Assuming that the pattern in Experiment 5a resulted from a syllable match

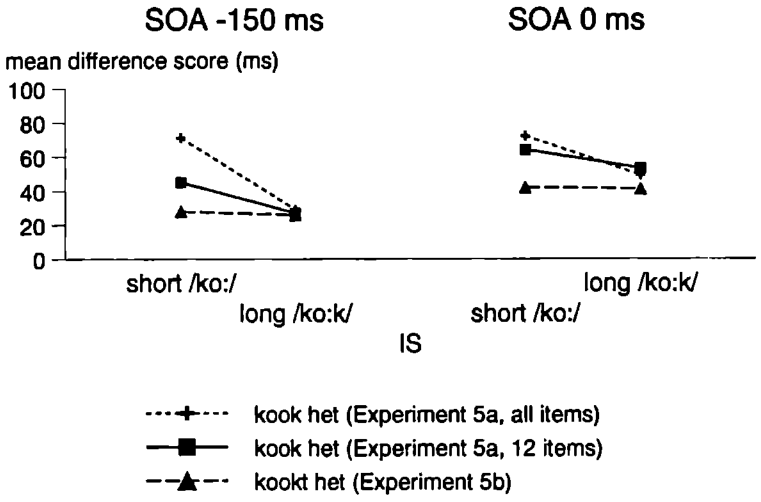


Figure 14. Mean Difference Scores for Short and Long Related IS from Pink Noise Baseline in Experiments 5a (Targets *kook het*, all Items), 5a (12 Items), and 5b (Targets *kookt het*, 12 Items)

effect, Experiment 5b should show the conversed pattern, since the targets in this study had a long first syllable, for instance, *kookt het* $(\text{ko:})_\sigma(\text{kət})_\sigma$. Related long IS should thus yield more facilitation than related short IS as compared to the pink noise baseline. This pattern, however, was not obtained. The amount of facilitation provided by short and long related IS did in fact not differ in Experiment 5b. Figure 14 shows the difference scores of related IS from pink noise for Experiments 5a and 5b. For Experiment 5a, the Figure shows the difference scores for all items and for

[9] The MSerrors used in the Dunnett's tests came from analyses that included the related and pink noise conditions. However, in Dunnett's tests calculated over all five treatments of IS condition, the difference of related long IS and pink noise was not significant at SOA -150 ms. For SOA 150 ms, the difference between related short IS and pink noise only reached .05-level over subjects and items.

the 12 items that had also been tested in Experiment 5b.

The difference scores of related IS as compared to the pink noise baseline in Experiments 5a and 5b were analyzed in an Anova that comprised experiment as a variable. Only those items of Experiment 5a were included in the analysis that had also been involved in Experiment 5b. It turned out that when only those 12 items of Experiment 5a were analyzed that had also been included in Experiment 5b, the pattern that had been obtained in the analysis over the whole item set of Experiment 5a changed (compare again Figure 14). Although short related IS still provided more facilitation than long related IS in the analysis of Experiment 5a with only 12 items, this difference was substantially reduced. With the 20 items that had been tested in Experiment 5a, the facilitation provided by short related IS had surpassed the facilitation provided by long related IS by 42 ms at SOA -150 ms, and by 23 ms at SOA 0 ms. The analysis with 12 items showed smaller values of 18 ms and 9 ms, respectively.

The subject Anova over the difference scores for related IS as compared to a pink noise baseline comprised the between-subjects variables 'experiment' (two levels: Experiment 5a, Experiment 5b), 'SOA' (two levels: -150 ms, 0 ms), and the within-subjects variable 'related length' (two levels: short and long). These variables were within-items variables in the item Anova. Both analyses revealed no significant effect or interaction. But one has to take into account that the power of the statistical analysis was reduced because of the smaller number of items that had been tested in Experiment 5b. The analysis over the whole item set of Experiment 5a suggested a syllable match effect. Perhaps an alternative control experiment could support this interpretation (see below).

Turning from Experiments 5a,b to the other experiments, the point for a syllable match effect is weakened. Related short IS provided more facilitation than related long IS as compared to a pink noise baseline for almost all conditions, irrespective of the targets' syllable structure (see Appendix 7.2). In Experiment 3, for instance, related short IS facilitated the production of past tense forms (e.g., *kookte*) more than related long IS at all SOAs, although the long IS correspond to the targets' first syllable.

An exception was again Experiment 2, which had tested single verbs in the infinitive and the past tense form, for example *koken* and *kookte*, see Figure 15. As compared to the pink noise baseline, short related IS provided more facilitation than long related IS for infinitives (69 ms and 57 ms, respectively), where the short IS correspond to the first target syllable. This pattern is predicted by the syllable match effect. The past tense forms, however, were facilitated equally strong by short (51 ms) and long (53 ms) related IS.

The Anova over related difference scores from pink noise revealed an interaction of length and verb form that just failed to reach significance in the analysis over both SOAs ($F_1(1;48) = 3.78$, $MSE = 729$, $p < .06$; $F_2(1;19) = 3.87$, $MSE = 565$, $p < .07$). Tests of simple effects showed that the interaction was not significant at SOA 150 ms, but again just failed to reach significance in the analysis over subjects at SOA 0 ms ($F_1(1;48) = 3.62$, $MSE = 729$, $p < .07$; $F_2(1;19) = 3.09$, $MSE = 689$, $p < .1$). Responses to infinitives where more facilitated by short IS than by long IS (85 ms for short IS and 67 ms for long ones), while short and long IS facilitated the responses to past tense forms equally well (58 ms and 61 ms, respectively). While the effect of IS length was not significant for the past tense forms at SOA 0 ms (all $F < 0$), it was significant over items and close to significance over subjects for the infinitive targets ($F_1(1;48) = 3.43$, $MSE = 1110$, $p < .07$; $F_2(1;19) = 5.79$, $MSE = 541$, $p < .03$).

The results for infinitive targets hence look like a syllable match effect. According to this effect, however, the facilitation provided by short and long IS should differ for the past tense forms: Long related IS should facilitate responses to past tense forms more than short related IS, because the first syllable of the past tense forms is long. This was not the case. The assumption that syllable match effect and segmental overlap effect co-occur does not predict the results, either. According to this assumption, the facilitation by long related IS should surpass the facilitation by short related IS in the past tense forms, while for infinitives, the facilitory effects of short and long related IS should not differ much. The opposite

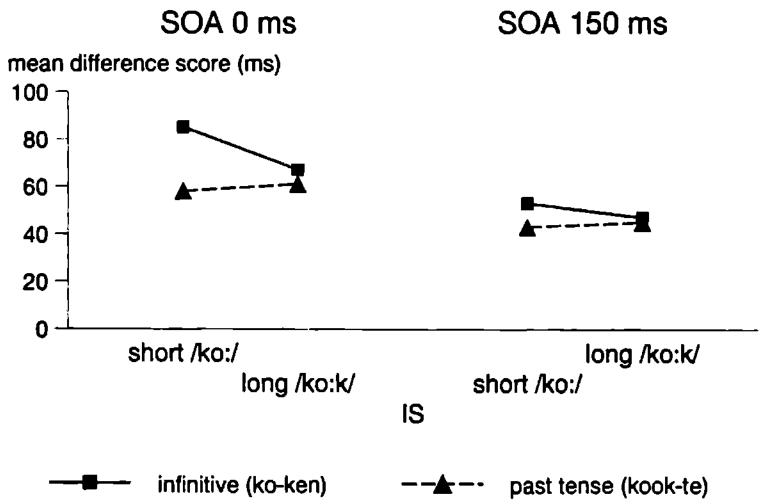


Figure 15. Mean Difference Scores for Short and Long Related IS from Pink Noise Baseline in Experiment 2 pattern was found for the difference scores of related IS from pink noise in Experiment 2. It cannot be caused by a syllable match effect alone, and neither by the co-occurrence of segmental overlap and syllable match effect. However, the syllable structure of the target word does seem to influence the obtained facilitative effects to a certain extent.

Replacing the unrelated baseline by pink noise hence does reveal effects that suggest that syllable structure may influence the facilitative effects to a certain extent, in combination with other variables. Only the pattern in Experiment 5a, could be interpreted as resulting from a pure syllable match effect: Related short IS yielded more facilication than related long IS as compared to the pink noise baseline for target words that had a short first syllable.

One should note that in this experiment, as in most other experiments, the same pattern that had been obtained for related IS also occurred for unrelated IS. Unrelated long IS yielded longer RTs than unrelated short ones and therefore slowed down responses more than short IS, as compared to the pink noise baseline. Looking again at Experiment 5a, Anovas over the difference scores of unrelated IS and pink noise revealed an effect of length, ($F_1(1;72) = 6.37$, $MSE = 1556$, $p < .02$; $F_2(1;19) = 4.74$, $MSE = 1689$, $p < .05$), and an interaction of SOA and length ($F_1(2;72) = 6.62$, $MSE = 1556$, $p < .01$; $F_2(2;38) = 6.1$, $MSE = 1355$, $p < .01$). Tests of simple effects showed that the difference in priming obtained for long and short unrelated IS was significant at all SOAs except at SOA 150 ms. Dunnett's tests showed that all RTs differed significantly from pink noise. When the pattern for related IS in this experiment is caused by a syllable match effect, the higher amount of interference that unrelated long IS showed in comparison to unrelated short ones has to be explained differently. It could be caused by the fact that more segments mismatch the target forms for the long than for the short unrelated IS. Following an alternative explanation, both patterns, in the related and unrelated IS conditions, could result from a main effect of length, which is caused by the physical length of the IS in the different conditions.

In sum, the RTs for related IS in Experiment 5a at first sight had suggested a syllable match effect when comparing them to pink noise as a baseline, since short IS, which corresponded to the targets' first syllable, facilitated reactions more than long IS. However, short IS led to faster RTs than long ones not only in the related, but also in the unrelated conditions of all experiments, including Experiment 5a. Assuming a syllable match effect for related IS, the pattern for unrelated IS has to be explained differently, by assuming a segmental mismatch effect. A main effect of length, on the other hand, according to which longer stimuli slow down RTs compared to shorter stimuli, covers the patterns of both relatedness conditions. A more important problem with the pink noise baseline is that the syllable match effect of Experiment 5a did not show up in the remaining experiments, except for the pattern in Experiment 2 that suggested a limited influence of this effect, but could not be expained by the syllable match effect alone. In the remaining experiments, short related IS provided more facilitation compared to pink noise for all target forms irrespective of the targets' syllable structure. Pink noise therefore does not provide an adequate baseline condition.

Instead of replacing the unrelated baseline with pink noise, one could also do without any baseline and consider only the RTs obtained for related IS. Compare, for instance, the RTs for related IS in Experiments 5a and 5b with targets of the type *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ and *kookt het* $(\text{ko:})_\sigma(\text{kət})_\sigma$, see Figure 16.

In the analysis over the whole data set of Experiment 5a, participants responded faster to the *kook het*-targets than to the *kookt het*-targets, when they heard short IS, whereas with long related IS, responses were faster to *kookt het*- than to *kook het*-targets. This suggests an influence of the target's syllable structure on response latencies. In the *kook het* - experiment, however, responses in the short and the long IS conditions did not differ in the *kookt het*-experiment, while they differed in the *kook het*-study. The assumption that segmental overlap and syllable match effect co-occur predicts the opposite pattern: When participants produce *kook*

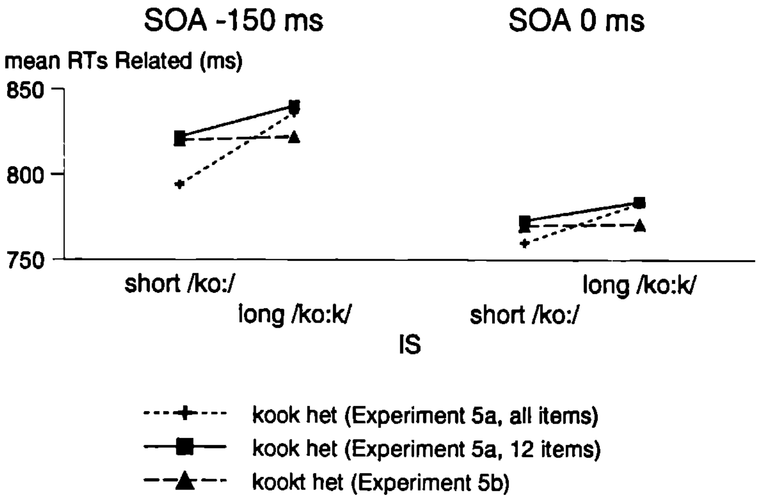


Figure 16. Mean Reaction Times for Short and Long Related IS in Experiments 5a (Targets *kook het*, all Items), 5a (12 Items), and 5b (Targets *kookt het*, 12 Items) *het*-forms, they should respond equally fast in the short and the long IS condition, because the targets' first syllable is short, and therefore RTs for short IS should be faster due to syllable match, while responses for long IS should be faster due to more segmental overlap. For the *kookt het* forms, on the other hand, responses for long IS should be faster due to both segmental overlap effect and syllable match effect and therefore should clearly differ from the responses for short IS.

The pattern obtained in Experiments 5a,b was in the opposite direction. It changed even more when only those 12 items of Experiment 5a were considered that had also contributed to Experiment 5b (compare again Figure 16). An Anova was conducted over the RTs obtained in the short and long related IS conditions of Experiments 5a,b. Only the 12 items were included that had contributed to both experiments. The subject Anova comprised the between-subjects variables 'experiment' (two levels: Experiment 5a, Experiment 5b) and SOA (two levels: -150 ms, 0 ms) and the within-subjects variable 'related length' (two levels: short and long). The item Anova comprised these three variables as within-items variables. The only effect that reached significance was the effect of SOA ($F_1(1;96) = 4.45$, $MSE = 29401, p < .04$; $F_1(1;11) = 55.96, MSE = 963, p < .01$). No other variable or interaction was significant. Again, one has to keep in mind that the number of items statistically analyzed was rather small. A further problem is the fact that Experiments 5a and 5b were two separate experiments. If exclusively the related IS conditions are taken into account, we lack a baseline to make sure that the results of the two experiments are comparable (but see the Summary and Discussion of Part 1 below for a possible solution of this problem).

Considering the results with the alternative pink noise baseline or without any baseline, one can conclude that the unrelated IS of corresponding length seems to be appropriate for analyzing the experiments that have been reported. If there is a main effect of length, in that long IS, both related and unrelated, yield slower responses than short ones, the unrelated baseline of corresponding length gets rid of it. Moreover, promising results were obtained with the unrelated baseline: The prime efficiency of short and long IS was found to significantly interact with the targets' syllable structure in Experiment 2, and also the late SOA of Experiment 4 showed a pattern that suggested a syllable match effect, although it did not reach significance.

Summary and Discussion of Part 1

The experiments reported in Part 1 of this thesis aimed at testing a prediction on the time course of syllabification in speech production, which had been formulated on the basis of Levelt's model of phonological encoding, see also the Hypotheses at the beginning of Part 1. The model involves an early stage, at which the segments of a word are made available, for instance /k, o:, k, ə, n/ for the target *koken*, and a late stage, at which segments are associated to a prosodic frame. The frame specifies the number of syllables of a word and its stress pattern. Syllables evolve from the association of the segments to the prosodic frame. Consequently, the model predicts a segmental overlap effect at an early point in time: When only the segments are available, a stimulus should reduce the time the participants need to produce the target the more efficiently, the more phonemes are shared by the stimulus and the target, as compared to an unrelated stimulus of corresponding length. The syllable structure of a target should play no role. At a later point in time, a syllable match effect should occur: Stimuli that correspond to the target's first syllable should be more efficient primes than stimuli that do not correspond to the target's first syllable.

In a series of experiments, participants produced targets that shared the initial phonemes, but differed in their syllable structure, for example, *kookte* $(\text{ko:k})_\sigma(\text{tə})_\sigma$ "cooked" or *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$ "cook (pl.) it", compared to *koken* $(\text{ko:})_\sigma(\text{kən})_\sigma$ "to cook" or *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ "cook (sg.) it". IS were presented at different SOAs with respect to the presentation of the cue that elicited the response. IS were either phonologically related to the target word, in that they shared phonemes, for instance, /ko:/ for the targets above, or unrelated, in that they did not share phonemes, for instance, /le:/ for the same targets. Furthermore, IS could be either short (e.g., /ko:/) or long (e.g., /ko:k/). The model's predictions concern the gain related IS provide compared to an unrelated IS of corresponding length to which I referred to as 'prime efficiency': At an earlier SOA, related long IS like /ko:k/ should be more efficient primes than related short ones like /ko:/ for all target forms, irrespective of their syllable structure, because they overlap in more segments with the targets than the short IS (= segmental overlap effect). At a later SOA, short IS should be more efficient than long ones for infinitives and encliticized forms of the type *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$, where they correspond to the targets' first syllable, whereas long IS should prime past tense forms and encliticized forms of the type *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$ more efficiently than short IS, since the targets' first syllable is long (= syllable match effect).

In a less strict formulation of the predictions, segmental overlap and syllable match effect co-occur. Under the assumption that the effects are additive and equally strong, short and long IS should be equally efficient in priming infinitives and encliticized forms of the type *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$, because short IS surpass long ones in priming due to the syllable match effect, while long IS should be more efficient than short ones due to the segmental overlap effect. Conversely, response latencies for past tense forms and encliticized forms of the type *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$ should be more efficiently primed by long IS, since these have more overlapping segments and match the targets' first syllable.

To summarize the results (see the Table in Appendix A7.1), **Experiment 1** showed that phonological priming can be obtained in a semantic-associate learning task with auditory syllabic IS and single word targets (e.g., *koken*, *kookte*). In this as well as in all following experiments, related IS led to faster responses than unrelated IS. **Experiment 2** included more targets than the first experiment, and a higher proportion of related IS. In this experiment, a main effect of length occurred: Participants responded slower when they heard long IS than when they heard short ones. This effect was accompanied by an interaction that was predicted for the case that segmental overlap and syllable match effect co-occurred: Long IS primed past tense forms better than short IS, while infinitives were primed by short and long IS equally well. Unexpectedly, the interaction was obtained at a rather early point (at SOA 0 ms), and disappeared at SOA 150 ms, while phonological facilitation was still huge at the later SOA. In **Experiment 3**, I tried to strengthen the effect by using IS that were closer to the phonetic level than the IS that had been used before. Instead of being spoken as syllables, these were cut out of the target verb forms. For instance, participants who named the infinitive *koken* ("to cook") heard IS like /ko:/ that had been cut out of *koken*, while participants who named the past tense form *kookte* ("cooked") heard /ko:/ cut out of *kookte*. However, the interaction obtained in Experiment 2 disappeared with these IS, while the phonological priming effect and the main effect of length were again significant. This suggests that the effect was located at an abstract phonological level. **Experiment 4** tested encliticized forms (e.g., *kook het*), including a low proportion of related IS. In addition to the phonological priming effect, a trend for a main effect of length was obtained. This showed that the paradigm could also be used for targets that contain more than a single word. **Experiments 5a,b** used the successful setup of Experiment 2 for encliticized forms of the type *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ "cook (sg.) it" (Experiment 5a) and *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$ "cook (pl.) it" (Experiment 5b). The analysis did not reveal an interaction similar to the one obtained in Experiment 2. **Experiment 6** was a control experiment that made sure that the obtained effects were independent from morphology, and **Experiment 7** excluded the possibility that they were caused by a lexical effect.

All experiments revealed a highly significant effect of phonological relatedness. Participants were faster when they heard related IS compared to unrelated counterparts of equal length. This shows that the SOA ranges (-150, 0, 150 ms) were appropriate for investigating phonological encoding, and furthermore demonstrates the stability of the priming effect.

Moreover, RTs were slower for long IS than for short IS. This main effect of IS length shows that the task is sensitive to small differences in the auditory IS. It was highly significant in Experiments 2 and 3, which both tested single verbs, and in Experiment 5a, which tested encliticized forms. In these experiments, the proportion of related IS had been 40%. Furthermore, the main effect of length gained significance over subjects in Experiment 4, which tested encliticized targets and included 12.5 % related IS. While the Morphological Control Experiment did not reveal a main effect of length, the interaction of relatedness and length was significant, showing that this experiment was also sensitive to the length variation. On the other hand, no significant main effect of length appeared in Experiment 1, which tested single verbs and included 10% related IS. The reason could be the small number of 10 target verbs in this experiment. Experiment 5b, which tested encliticized targets of the type *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$, lacked a main effect of length, too. But this experient also included only 12 targets. Furthermore, participants seemed to find the targets more difficult than the encliticized forms of the type *kook het* that had been used in Experiment 5a.

In the Morphological Control Experiment, long IS (e.g., /bo:t/) primed morphologically simple target words like *boter* ("butter") and complex targets like *boten* ("boats") equally well, although they corresponded to the complex targets' stem morpheme (e.g., *boot* "boat"), but not to a morphological unit in the simple words. This showed that the obtained effects in Experiments 1 to 5 were not caused or hidden by morphological effects. In the Lexical Control Experiment, word and pseudoword IS did not differ in prime efficiency. This excluded potential lexical effects that could have been arisen because most long IS were Dutch words, whereas most short IS were not. This variable hence did not influence the results obtained in Experiments 1 to 6.

In the final section of Part 1, I discussed whether the experimental results could be reinterpreted using a different baseline than unrelated IS of corresponding length. The results of Experiment 5a suggested to compare the RTs for related IS with pink noise as a baseline, because short IS provided more facilitation than long IS in Experiment 5a. In this experiment, they corresponded to the targets' first syllable, like the /ko:/ in *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$. Unfortunately, Experiment 5b, which tested targets whose first syllable corresponded to the long IS, did not reveal a similar pattern. According to the syllable match effect, long IS should have provided more facilitation than short ones, but the results showed that short and long IS provided an equal amount of facilitation. However, the results of Experiment 5b should not be overvalued, since the targets in this experiment seemed to be more difficult than in Experiment 5a. A further argument against the pink noise baseline is that in other experiments, either related short IS provided more facilitation than related long ones as compared to pink noise for all target forms, irrespective of the targets' syllable structure (Experiment 3), or the pattern could be explained by neither the syllable match effect alone, nor by the co-occurrence of segmental overlap and syllable match effect (Experiment 2).

The unrelated baseline of corresponding length that has been used throughout Part 1 seems to be an appropriate one. With the unrelated baseline, a promising result had been obtained. The pattern that was expected under the strict version of the predictions for the time course of syllabification - an early segmental overlap and a late syllable match effect - never showed up in the experiments. But Experiment 2 showed an interaction that could be interpreted as resulting from the co-occurrence of segmental overlap and syllable match effect. While infinitives like *koken* $(\text{ko:})_\sigma(\text{kən})_\sigma$ were primed equally well by short and long IS, long IS were more efficient than short IS in priming past tense forms, which have a long first syllable, like *kookte* $(\text{ko:k})_\sigma(\text{tə})_\sigma$. Unexpectedly, the interaction occurred at SOA 0 ms and disappeared later, while phonological facilitation was still obtained at the later SOA. In Experiment 4, there was a strong tendency for short IS priming encliticized forms of the type *kook het* $(\text{ko:k})_\sigma(\text{tə})_\sigma$ more efficiently than long IS, which did not match the targets' first syllable. This pattern, which looks like a syllable match effect, occurred as predicted at a late SOA.

In Experiment 5a, short and long IS had a similar prime efficiency for encliticized forms like *kook het*. This had been predicted for the case that segmental overlap and syllable match effect co-occurred. The control Experiment 5b with forms like *kookt het*, however, did not show the predicted pattern: Instead of a higher prime efficiency for long IS, there was a trend for short IS being more efficient than long ones at an early point in time, while at a later SOA the efficiency of short and long IS did not differ. Whereas participants in Experiment 5a named singular imperatives like *kook het* ("cook it") without difficulties, they had to produce plural imperatives in Experiment 5b (*kookt het*, "cook (pl.) it") and reported that they found those rather difficult to produce. One could try to avoid this problem and moreover to increase the obtained effects in a study that is changed in three respects.

The first change concerns the target forms. An alternative encliticized form that has a long first syllable and can be used as a counterpart for the imperative *kook het*, which has a short first syllable, is the question *kookt het?* ("does it cook?"). This question does not differ from the plural infinitive forms that had been used in Experiment 5b phonologically, but may be easier pragmatically. Since the imperatives (*kook het!*) and the questions (*kookt het?*) are compared, they should include the same target verbs and the same semantically associated cue nouns. Therefore, we need target verbs for which the cue noun can either be interpreted as an object (for the imperatives: *soep* - *kook het*, "soup - cook it") or as a subject (for the questions: *soep* - *kookt het?*, "soup - does it cook?"). Perhaps these new targets would yield an effect similar to the one obtained in Experiment 2: Long IS should be more effcetive primes for the questions than short ones, because they match with the questions' first syllable.

A second change could be to include only phonologically related IS conditions. The experiment would test both the imperative targets and the question targets (e.g., *kook het!* and *kookt het?*), in a related short and related long IS condition (e.g., /ko:/, /ko:k/). According to the segmental overlap effect, participants should name both targets forms faster when they hear a long IS like /ko:k/ than when they hear a short IS like /ko:/. According to the syllable match effect, participants should name the imperatives faster when they hear the short, and name the questions faster when they hear the long IS.

A third change results from the preceding two suggestions. Because the experiment will be shorter, since no unrelated IS conditions have to be included, one could vary the SOA within subjects. Since one subject produces all targets at all SOAs and in all IS conditions, a baseline is not necessarily needed.[10]

In sum, whether the first prediction that was tested in Part 1, namely that syllables are produced late during phonological encoding, is correct, is still an open question. The results of Experiment 2, where the different prime efficiency for short and long IS interacted with the syllable structure of the target forms, was promising and showed that the syllables are relevant during phonological encoding. Further research has to investigate the time course of their production more closely.

---

[10] In the meantime, an experiment along these lines has been done. Sixteen verbs were produced either as an imperative with a short first syllable (e.g., *kook het!*), or as a question with a long first syllable (*kookt het?*). Related IS were presented at SOA 150 ms which corresponded either to the short or to the long first syllable. Both short and long IS speeded up RTs as compared to a pink noise baseline. There was thus a phonological priming effect. Furthermore, participants responded faster when they heard long than when they heard short related IS. This effect is in the opposite direction as the main effect of length obtained in the previous experiments, where participants responded faster to short than to long IS, irrespective of phonological relatedness. This length effect, however, was obtained both for targets with short and long first syllables. Thus, again there was no effect of syllable match.

```
st
  und
     en
  und
st
  und
     en
  und
st
  und
     en
  und
st
  und
     en
```

(ernst jandl)

PART 2: THE LEVELS OF SYLLABIFICATION

Two sets of experiments investigated how many levels of syllabification are involved when speakers produce encliticized forms. While in Levelt's model of phonological encoding the postlexical syllables are produced immediately, an alternative account contains two levels of syllabification, where a lexical level precedes a postlexical one. The phenomenon of syllable-final devoicing in encliticized and other contexts was investigated to find out which account is correct.[11]

Final Devoicing in Encliticized Forms: Phenomenon, Predictions, Paradigm

Final Devoicing in Dutch

In Levelt's model of phonological encoding, lexical forms are not syllabified. The only syllables that are built are the surface syllables, which do not always coincide with lexical boundaries in connected speech. In Dutch encliticized forms,

[11]Parts of the results reported in this section are in press in A. Dainora, R. Hemphill, B. Luka, B. Need, & S. Pargman (Eds.), *CLS 31-II: Papers from the Parasession on Clitics*, Chicago: Chicago Linguistic Society.

for instance, an enclitic will prefer to take a coda element of the preceding word as onset, to avoid a schwa-initial syllable. This was shown in the Introduction as shown in (3), and I repeat it for convenience in (4):

(4)	$(\text{drɪŋ})_\sigma(\text{kət})_\sigma$	*drink het*	"drink it"
	$(\text{daŋ})_\sigma(\text{kəm})_\sigma$	*dank hem*	"thank him"
	$(\text{bo:})_\sigma(\text{tən})_\sigma$	*boot en*	"boat and"

How can we test the model's claim that the abstract lexical syllables are never produced during speaking? Dutch, as many other languages, has constraints for codas that do not hold for onsets. If the obstruent that surfaces in onset position of the clitic's syllable, like the /k/ in $(\text{drɪŋ})_\sigma(\text{kət})_\sigma$, nevertheless shows effects of coda constraints, this would imply that the obstruent has been in a coda position at an earlier stage of processing before it became an onset in surface syllable structure.

One such constraint for Dutch codas is that only voiceless obstruents occur in syllable codas, while both voiced and voiceless obstruents occur in onsets.

(5)	a.	*nies*	$(\text{ni:s})_\sigma$	"sneeze (1st pers. sg)"
	b.	*niezen*	$(\text{ni:})_\sigma(\text{zən})_\sigma$	"to sneeze"
	c.	*rood*	$(\text{ro:t})_\sigma$	"red"
	d.	*rode*	$(\text{ro:})_\sigma(\text{də})_\sigma$	"red (inflected)"
	e.	*hand*	$(\text{hant})_\sigma$	"hand"
	f.	*handig*	$(\text{han})_\sigma(\text{dəx})_\sigma$	"handy"

As shown in (5), voicing is maintained in inflected (5b, 5d) or derived (5f) forms where the suffix starts with a vowel, indicating that it is located after suffixation has taken place. One could argue that final devoicing applies at the end of a morpheme that is followed by a vowel-initial suffix. Thus, it will apply in *schrijfster* ("female writer"), which is derived from the verbal stem *schrijv-*, but not in the form *schrijver* ("male writer"), where the suffix is *-er*. However, the distinction of suffixes into two classes (V-initial versus C-initial) is ad hoc, since those two groups do not serve for defining any other processes. Moreover, the pronunciation of acronyms in Dutch gives evidence that final devoicing takes place syllable-finally (Booij, 1995). For instance, *ABVA*, the acronym for *Algemene Bond van Ambtenaren* ("General Union of Civil Servants") becomes $(\alpha\text{p})_\sigma(\text{fa})_\sigma$. Finally, syllable-final devoicing is productive. This becomes obvious when Dutch speakers speak languages that allow voiced obstruents in codas (like English) - final devoicing cannot easily be 'switched off'.

Final devoicing has been seen traditionally as a rule that changes the voice feature associated with an underlyingly voiced obstruent into a voiceless feature.

More recent approaches in phonology have replaced the notion of phonological rules by wellformedness constraints. A constraint-based analysis of syllable-final devoicing has been provided by Lombardi (e.g., Lombardi, 1995). She argued that syllable-final devoicing could be analyzed as a positive wellformedness constraint on the presence of phonetic feature specifications. Instead of a rule delinking features in certain positions, like the voice feature in coda position, a wellformedness constraint allows features in certain environments.

Lombardi's analysis bases on two assumptions. First, she interprets [voice] as a privative feature. This means that a segment is voiced when it is associated with the feature [voice], and if it is not associated to the feature [voice], the segment will take the unmarked value and be voiceless. A second assumption Lombardi makes is that the voice feature is dominated by a laryngeal node in feature geometry. The term feature geometry refers to a hierarchical structure that phonologists have introduced to capture groupings of features (see, for instance, Clements, 1985). For instance, vowels can assimilate to a neighbouring vowel with respect to the place features [back], [high], and [low], which all refer to the position of the back of the tongue (the dorsum), and are therefore grouped under the articulator feature [dorsal]. Vowels can also assimilate to the nasal specification of a neighbouring sound. The feature [nasal] is dominated by the articulator feature [soft palate]. But vowels do not assimilate to a neighboring sound, for instance, with respect to [nasal] and [back], but not [high] and [low]. The features [back, high, low] group together, and this is expressed in the hierarchical feature tree.

The feature [voice] groups, together with features that involve the position of the glottis, under a [laryngeal] articulatory feature node. Lombardi argues that the absence of voiced obstruents in syllable codas is caused by a wellformedness constraint that allows an laryngeal articulator feature node only for onsets. In coda position, the laryngeal node is not allowed. Since the feature [voice] is grouped under the laryngeal node, the feature [voice] will not be present in codas, either. This leaves a consonant without voice specifications. According to Lombardi's first assumption mentioned above, this consonant will take the unmarked value and be voiceless.

Lombardi's account explains why neutralization of laryngeal features never produces only voiced stops, while a rule-based account has to stipulate this (for a further evaluation of Lombardi's proposal see Kenstowicz, 1994, p. 493 ff). However, whether devoicing results from a rule or from a wellformedness constraint is not crucial for the present argument. I will use the traditional term 'rule', but could as well have used 'constraint'. Both, rule and constraint, are defined on the basis of syllabic structure. The crucial question for the present investigation is the level of syllable structure that is relevant for final devoicing in encliticized forms.

Final Devoicing in Encliticized Forms: Predictions

For encliticized forms, it has been claimed that final devoicing applies on the single lexical items, before postlexical resyllabification. Following this account, word-final obstruents are devoiced although they surface in onset position of the following syllable (see e.g. Kooij, 1980; Booij, 1996)[12], as shown in (6a).

(6a)	*bind het*	$(\text{bɪn})_\sigma(\text{tət})_\sigma$	"bind it"
	vriend en	$(\text{vri:n})_\sigma(\text{tən})_\sigma$	"friend and"
(6b)	*bind het*	$(\text{bɪn})_\sigma(\text{dət})_\sigma$	"bind it"
	vriend en	$(\text{vri:n})_\sigma(\text{dən})_\sigma$	"friend and"

The inflected forms show that the obstruents are underlyingly voiced: *binden* $(\text{bɪn})_\sigma(\text{dən})_\sigma$ "to bind", *vrienden* $(\text{vri:n})_\sigma$ $(\text{dən})_\sigma$ "friends". Levelt's model, on the other hand, predicts the forms in (6b). Since the surface syllable structure is the only syllabic structure created during the production process, the first word's final obstruent is never in syllable-final position, where it could be devoiced. The experiments reported below were run to test which account makes the right predictions.

Experimental Studies on Final Devoicing

Most experimental studies of final devoicing investigated syllable-final devoicing in German and word-final devoicing in Polish. They addressed the question whether the phonological voice contrast in obstruents is completely or only

---

[12] There has been some discussion about final devoicing in cliticized forms (see Berendsen 1986; Booij & Rubach, 1987). This discussion focussed on some combinations of modal verbs and clitics, for which final devoicing does not seem to be obligatory, for instance, *heb ik* ("have I"), for which both forms are attested: /hɛ.bɪk/ and /hɛ.pɪk/. Some authors account for this by different prosodic structures for the two forms. While clitics in the voiceless variant attach to the prosodic phrase, they adjoin to the prosodic word in the voiced variants and survive due to resyllabification within the prosodic word (Berendsen, 1986). Booij (1995, 1996), on the other hand, assumes that the voiced variant is stored as one unit in the lexicon due to the high frequency of occurrence of the modal verbs. Devoicing only occurs when the encliticized forms are constructed instead of retrieved as a whole from the lexical store. This discussion is still open, but not crucial for the present argument, since only full verbs were tested.

partially neutralized in final position phonetically. The studies were either perception experiments or included acoustic measurements, or combined both approaches. The results were contradictory.

Participants in perception studies hear minimal pairs of words that only differ in the underlying voice quality of their final consonant. An example is the Dutch pair *raad* ("commission") and *raat* ("honeycomb"). In the plural forms of these words, the stops clearly differ. They are in onset position and the voiced stop remains voiced: *raden* $(\text{ra:})_\sigma(\text{dən})_\sigma$ - *raten* $(\text{ra:})_\sigma(\text{tən})_\sigma$. In the syllable-final position of the singular forms, however, the voiced stop devoices. Some perception studies found complete neutralization of the voice contrast in word-final position, like for Polish Jassem and Richter (1989), or for German Kahlen-Halstenbach (1990). Participants could not distinguish members of minimal pairs like *raat* and *raad*. Port and O'Dell (1985), on the other hand, concluded from a 59% correct performance in a forced choice task on German minimal pairs that syllable-final devoicing does not completely neutralize the voice contrast. Slowiaczek and Szymanska (1989) concluded the same for Polish.

Among the studies that included acoustic measurements, some found complete neutralization. These include Jassem and Richter (1989) for Polish, Fourakis and Iverson (1984) for German. Others found a difference between underlyingly voiced and voiceless obstruents, like Slowiaczek and Dinssen (1985), who attested a significantly longer preceding vowel duration for underlyingly voiced compared to voiceless stops in Polish, Port and O'Dell (1985), who report a small but significant similar effect for German, and Port and Crawford (1989), who used several tasks to elicit German utterances that contained minimal pairs that only differed in the voicing of their final consonant. They found incomplete neutralization in tasks where the target words were contrastively stressed and in a dictation task. When participants had to read or repeat sentences with normal intonation, however, the correct categorization in a discriminant analysis dropped to 55%. Since only three minimal pairs were tested which moreover included words of different word classes, the impact of their results is limited. Others found both complete and incomplete neutralization depending on the context. Dinssen and Charles-Luce (1984), for instance, found that the closure duration and the duration of the preceding vowel differed for voiced and voiceless Catalan stops depending on the phonetic context and the particular stop consonant. Charles-Luce (1985) found the acoustic parameters to vary with phonetic and syntactic context in German, and Piroth, Schiefer, Janker, and Johne (1991) obtained complete neutralization in German when the stop preceded a morpheme boundary, but incomplete neutralization across word boundaries or utterance-finally (for summaries see also Dinssen, 1984; Brockhaus, 1989). For Dutch, Jongman, Sereno, Raaimakers, and Lahiri (1992) reported a complete neutralization of the voice contrast.

The studies run so far are not without problems. In most of them the utterances were read from lists. This might encourage artificial and hypercorrect pronunciations. Moreover it has been seen as problematic for languages like German, which code the underlying voice contrast orthographically, also in neutralizing positions. Some studies therefore tried different paradigms, like verb conjugation (Fourakis & Iverson, 1984), where participants were told an infinitive and had to produce the preterite form which ended in the obstruent that was devoiced, like German *meiden* - *mied* ("to avoid - avoided"). This study revealed a complete neutralization of the voicing contrast. Other researchers asked participants to repeat words from memory that had been read to them by an assistant (Port & Crawford, 1989). However, since the assistant probably knew the aim of the experiment, the pronunciation was perhaps not naive anymore.

The minimal pairs, at least for German, often included very low-frequent words, like *Alb* ("elf", also name of a mountain region) - *Alp* (demon believed to cause nightmares, also "alpine pasture"), *schrak* (old preterite of "frighten") - *schrag* (no German word at all). Port and O'Dell (1985), where the examples are taken from, gave definitions along with the word pairs to the participants in the production and perception task to make sure that they knew all words. A further problem with this study was that four of the ten participants in the perception experiment had already participated in the production task. Furthermore, the set of minimal pairs sometimes included only three (Port & Crawford, 1989) or four (Charles-Luce, 1985) pairs. Another problem was that the studies ran mainly in American laboratories and some participants had been living in the English-speaking environment for several years and were very skilled speakers of English, which is a language that maintains the voice contrast in final position.

The present experiments try to avoid some of the problems. Instead of reading isolated words or sentences from a list, speakers produced minimal pairs in context in a delayed repetition task. Furthermore, Dutch has a higher number of minimal pairs than German. The members of German minimal pairs mostly belonged to different word classes, like *bunt* ("colorful") - *Bund* ("bond"), or *seit* ("since") - *seid* ("you/pl. are"). In contrast to that, the stimuli of the present experiments were all minimal pairs of which both members were nouns. This avoids the inclusion of another variable 'word class' in the experiment.

Paradigm

The empirical investigation of final devoicing in encliticized forms consisted of two sets of two experiments. Each set contained a production and a perception experiment. The aim was to obtain information about the voice quality of the stop consonant in forms like *raad en* ("commission and"), where the weak form of the conjunction *en* /ən/ encliticizes to the preceding noun and the stop surfaces in onset position of the second syllable.

The stimuli consisted of minimal noun pairs that only differed in the underlying voice quality of their final obstruent, like *raad* ("commission") and *raat* ("honeycomb"). In the singular forms, this difference is neutralized by final devoicing. In the plural forms, however, the voice quality of the stops is maintained, since they are in onset position of the second syllable, as shown in (8).

(8)	singular			plural			
	raad	(ra:t)$_\sigma$	"commission"	*raden*	(ra:)$_\sigma$(dən)$_\sigma$	"commissions"
	raat	(ra:t)$_\sigma$	"honeycomb"	*raten*	(ra:)$_\sigma$(tən)$_\sigma$	"honeycombs"

In the production experiment, participants produced sentences that contained the minimal pairs in different contexts. In the perception experiment, different participants heard the sentences that had been produced in the first experiment and had to perform a rating task on the voice quality of the critical obstruent.

The noun could occur in four different contexts. In a 'final' context condition, the critical obstruent occurred sentence finally. In a 'nasal' context condition, the obstruent was followed by a nasal consonant. In both contexts, final devoicing should apply obligatorily. In the nasal condition, I expected in addition a voice assimilation effect from the nasal on the preceding stop. Therefore, the stop should be more voiced than in the final context in both, voiced and voiceless targets. In a 'plural' condition, the minimal pairs should be clearly distinguishable. In a 'clitic' context condition, the nouns were followed by a schwa-initial function word. The model of phonological encoding predicts for this condition that participants in the perception experiment should be able to distinguish nouns with underlyingly voiced from those with voiceless stops. This is because the stop is never in syllable-final position, where it could be devoiced. Following a theory that includes resyllabification, on the other hand, participants should not be able to distinguish between underlyingly voiced and voiceless nouns, since final devoicing applies on the individual lexical words, preceding resyllabification in the clitic context.

Experiment 8a: The First Production Experiment

Method

Participants. Twenty students (10 men and 10 women) took part in the production experiment. Every student participated in only one of the final devoicing experiments.

Stimuli. The stimuli were minimal noun pairs that only differed in the underlying voice quality of the last obstruent (e.g. *pond* "pound" - *pont* "ferry"). Ten pairs ended in the apical obstruents /t/ and /d/, and two pairs ended in the labials /b/ and /p/. This is almost the whole set of monosyllabic minimal noun pairs varying in the voice quality of their final stops in Dutch (= variable underlying consonant with two levels: voiced and voiceless).[13]

There are no minimal pairs ending in voiced or voiceless fricatives in Dutch, because voiced and voiceless fricatives are accompanied by different vowels: Vowels preceding the underlyingly voiceless fricatives are lax, while vowels preceding the voiced fricatives are tense (e.g. /ka:zən/ "cheeses" vs. /kɑsən/ "cash registers") . In addition, many speakers, in particular from the Western part of the Netherlands, neutralize the underlying voicing distinction between /f/ and /v/, and also the distinction between /x/ and /ɣ/, in word-initial position (Booij, 1995). There seems to be a general tendency for the voiced fricatives to disappear in Dutch: Slis and van Heugten (1989) found voiced realizations of /z,v/ in only about 25% of the cases for speakers from the South-Eastern part of the Netherlands - a region in which the voicing distinction is said to be completely preserved - and in only 20% of the cases they found voicing for speakers from the western region. Furthermore, fricatives between sonorants are voiced by a general fricative voicing rule. Thus, both underlyingly voiced and voiceless fricatives would be either voiceless (due to the decrease of voiced fricatives in Dutch) or voiced (due to the general voicing rule). Therefore, only stops were included.

The minimal pairs in the experiment occurred in three different contexts. In a final context condition, the noun appeared sentence-finally. In this absolute sentence-final position it has to be realized as a voiceless stop. In a nasal context the noun was followed by a nasal consonant that does not allow resyllabification. If the critical final obstruent was to form an onset of the following syllable together with the consonant following it, this would result in an onset cluster of obstruent and

---

[13]One additional minimal pair was found afterwards. It was included in the second set of experiments.

nasal that is forbidden in Dutch, for example, */dn-/, */pn-/. However, the nasal should induce phonetic regressive voice assimilation on the preceding obstruent. This results in three levels of the variable context: final, nasal, clitic. In a clitic context the target noun was followed by the schwa-initial function word *en* ("and"). This is the context for which the two accounts make different predictions.

An alternative enclitic function word might have been the pronoun *het* $(\mathrm{ət})_\sigma$ "it". Although verbs are probably the most common host for the weak form of *het* (e.g., *kook het* "cook it"), sentences could be constructed where it attaches to a preceding noun, for instance in subclauses like *Zorg ervoor dat de raad het was* ("Take care that the commission it was"). Problematically, however, the pronoun ends in a voiceless [t] and the voice contrast under investigation was [t/d] in almost all minimal pairs. One might argue that if an underlyingly voiced stop was perceived as voiceless, this actually occurred because of the neighbouring identical voiceless stop of the pronoun. Therefore, the conjunction was chosen.

Several means were introduced to distract participants' attention from the fact that phonology was important and that there were minimal pairs in the materials. Participants were told that they participated in a syntax memorization task. The carrier sentences were slightly varied syntactically. There were six different carrier sentences (see 9).

(9)	*Pien zegt "Ik zie een ... ".*	"Pien says 'I see a ... '."
	Pien zegt "Ik zag een ... ".	"Pien says 'I saw a ... '."
	Pien zegt "Er is een ... ".	"Pien says 'There is a ... '."
	Pien zegt "Er was een ... ".	"Pien says 'There was a ... '."
	Pien zegt "Ik heb een ... ".	"Pien says 'I have a ... '."
	Pien zegt "Ik had een ... ".	"Pien says 'I had a ... '."

Since it is not easy to get speakers to produce encliticized forms in an experimental situation, the carriers were rather long and had to be produced in a small amount of time. This time pressure made participants use encliticized forms (but see below).

The dots in (10) stand for a sequence of the minimal pair noun and a second noun, for instance, in the nasal context: *een raad naast een akker* ("an advice near a field"). A native speaker of Dutch checked that the two objects in the sentence were not semantically related and that all sentences were of similar "nonsenseness". The phonological restrictions on the materials did not allow for the construction of semantically "normal" sentences. The second noun in the sentence (*akker* "field" in the example above) was different for every member of the minimal pairs to distract the participants' attention from the fact that the materials included minimal pairs of target words. It was always a vowel-initial trochee. The reason for controlling the stress pattern of the second noun was its potential influence on the rhythmic phrasing of the sentence. As a consequence of these differences in phrasing, the realization of the final obstruent of the first noun could vary. The nucleus of the second syllable of the second noun was always a schwa. Two members of a minimal pair were combined with second nouns starting with the same vowel, as e.g. in *raad - akker* ("commission - field") and *raat - akte* ("honeycomb - file") or in *lab - otter* ("lab - otter") and *lap - orgel* ("cloth - organ"). Table 16 shows an example for the different context and voicing conditions.

context	voicing	target sentence
clitic	voiced	*Pien zegt "Ik zie een* ***raad en*** *een akker".*
	voiceless	*Pien zegt "Ik zie een* ***raat en*** *een akte".*
nasal	voiced	*Pien zegt "Ik zie een* ***raad naast*** *een akker".*
	voiceless	*Pien zegt "Ik zie een* ***raat naast*** *een akte".*
final	voiced	*Pien zegt "Ik zie een akker en een* ***raad****".*
	voiceless	*Pien zegt "Ik zie een akte en een* ***raat****".*

Table 16. Sentences to be Produced for the Minimal Pair *raad - raat*

In addition to the 12 experimental pairs, an equal number of fillers was included. As the experimental stimuli, they differed in one segment. Four pairs differed in onset, four in nucleus, and four in coda position (like *kan* "pitcher" - *pan* "pan"; *rek* "rack" - *rok* "skirt"; *been* "leg" - *beer* "bear"). As the members of the minimal pairs, the filler nouns were low frequent, but known words. They were distributed over the six carrier sentence types in such a way that each carrier type occurred with two pairs that were manipulated at a different syllable position, for example, with *schil/schim* (coda) and *boon/troon* (onset).

The 24 filler words were combined with six vowel-initial monosyllables, six consonant-initial monosyllables, and six pairs of bisyllabic initially stressed nouns starting with a consonant. The second nouns combined with the members of a filler pair had some segmental overlap, like the second nouns for members of a minimal pair. The bisyllabic words were included to introduce a different metrical pattern to prevent that the participants adapting a fixed rhythm. The stimuli can be found in Appendix A8.1.

Apparatus. The experiment was run with the NESU setup. Participants were seated in a soundproof booth. Their responses were recorded on one channel of a DAT-tape using a SONY 55 ES recorder and a Sennheiser ME 80 microphone. Two tones of different frequency on the second channel of the DAT-tape marked the presentation of the prompt on the participant's Samotron SC-428 TXL monitor and the speaker's voice onset. With the tones it could be checked whether the participants really waited for the prompt before they produced the sentence and did not simply read it off the monitor. The responses were registered by a voice key. The experimenter could see the visual information given to the participant, the reaction times, and the trial sequence on two control screens.

Design. Since the purpose of this experiment was to produce materials for the perception study, there was no full factorial design. The 144 sentences (12 target minimal pairs and 12 filler pairs x 2 underlying voice x 3 phonological contexts) were randomized in two blocks. Every item occurred twice, once in the first, and once in the second block. Each block was randomized differently. Not more than three experimental trials followed each other, and the same phonological context condition occurred maximally three times in a row. Furthermore, at least three trials intervened between two members of a target minimal pair. Five sets of the two blocks were presented to five different participant groups.

Procedure. A delayed repetition task similar to the one used by Ferreira (1991) was adopted to elicit participants' responses. Participants sat in front of a computer screen. A sentence appeared on the screen. Participants had to memorize the sentence and to produce it in reaction to a visual prompt on the screen. In contrast to Ferreira's study, the presentation of the sentences was not paced by the participants, but fixed. The delayed repetition task should delimit the influence of orthography, which is desirable, because the voicing contrast is coded in the spelling of Dutch.

Participants were tested individually. They were told that sentences would be presented on the monitor for a short period of time and their task was to memorize those sentences and to reproduce them as quickly and accurately as possible as soon as a visual cue occurred on the screen. They were told that the goal of the experiment was to investigate how easily different syntactic constructions can be memorized. Furthermore they were informed that the sentences were semantically odd and were quite similar to make the task more difficult. The session started with the word *Attentie* ("attention") appearing on the screen for 3500 ms. Then, the first block started in which all sentences appeared once, preceded by six practice sentences, which were two sentences of every sentence type. The word *Pauze* ("pause") appeared on the screen for three seconds after the last sentence of the first block, and there was a pause of about two to five minutes (the length was determined by the participant). Then, after another attention sign, the same six practice sentences occurred, followed by the second exposure of the items, which ended with *Einde* ("end") appearing on the screen. A single trial had the following structure: The sentence appeared in the center of the screen for 1500 ms. After a random pause of 500 to1000 ms, the prompt "*?????*" appeared in the center of the screen for 500 ms. The onset of the speakers' response was registered by the computer. The following trial began 2000 ms after the prompt had appeared on the screen.

The experiment took about 35 minutes. After the experiment, participants had to fill in a question form about their place of birth and the time they had been living at different regions in the Netherlands.

Experiment 8b: The First Perception Experiment

In the perception experiment, participants had to decide which member of a minimal pair had been produced in the production experiment by means of a 5-point-scale.

Method

Participants. Twenty students (10 men, 10 women) participated in the experiment. Five men and five women came from the South of the Netherlands, and the others from the rest of the country.

Stimuli. Participants heard sentences, for example, *Pien zegt "Ik zie een raad en een oom"*, produced by four speakers from the production experiment. Slis and van Heugten (1985) showed that female speakers produce less voicing in fricatives than male speakers in Dutch. Furthermore, voice assimilation patterns in consonant clusters differ for men and women (Slis, 1986). Therefore, two male and two female speakers were selected. One male and one female speaker were from the Southern part of the Netherlands, one male and one female speaker came from the Western part of the Netherlands. All of them had produced differently randomized lists in the perception experiment. They had no speaking disorder, used no strange intonational patterns, and had an intelligible voice. Furthermore they had produced every sentence without an error at least once. Of the two repetitions that had been produced of each experimental item in the production experiment, the first one was chosen, except when it had been incorrect or interrupted, or spoken hesitantly.

Five of the twelve minimal pairs of the production experiment had to be excluded, because one member was preceded by an article, while the other member was a mass noun and must not be preceded by an article, like *Er was een noot en* ... ("There was a nut and ...") versus *Er was nood en ...* ("There was poverty and ..."). So listeners could definitely decide which member of the minimal pair they heard based on the presence or absence of a determiner.

The second noun that occurred in the sentence - in addition to the target noun - had been different for the two members of a minimal pair in the production experiment to distract the speaker's attention from the minimal pair design. For the perception experiment, however, the context had to be identical. Therefore, in all sentences, the second noun was replaced by the noun *oom* ("uncle"), which was taken from a filler sentence of the production experiment and was not semantically related to any of the target nouns. Participants heard the minimal pairs, for instance, *raad* and *raat*, in three contexts, either at the end of an utterance (e.g., *Pien zegt "Ik zie een oom en een raad/t"*), or followed by a nasal (e.g., *Pien zegt "Ik zie een raad/t naast een oom"*), or followed by a schwa-initial enclitic pronoun (e.g., *Pien zegt "Ik zie een raad/t en een oom"*). Every participant heard seven minimal pairs in three contexts spoken by four speakers, yielding a total of 168 sentences.

In addition, twelve practice items were included. They consisted of three pairs taken from filler sentences: *mop - mot, schim - schil, kas - kaf.* Three productions from each of the four speakers appeared in the practice block, and each of the three contexts appeared four times. A member of a pair stood once in left and once in right position on the screen, and the correct answer for a pair was sometimes left and sometimes right. The materials are given in Appendix A8.2.

Apparatus. The experiment was again run with NESU. Participants were seated in a soundproof booth in front of a Samotron SC-428 TXL monitor. On the table before them were a pushbutton device and a computer keyboard connected to a Hermac AT computer. The auditory stimuli were digitized with a sample frequency of 20 kHz and presented over Sennheiser HD 250 headphones. The experimentator could see the participants' visual input and the response keys pressed by the participant on two control screens. The response keys were automatically written to a result file. Furthermore, the experimenter could hear the auditory stimuli presented to the participants.

Design. The design included three types of context (clitic, final, and nasal), two types of underlying voicing (voiced and voiceless), and four speakers.

Procedure. Participants listened to the sentences that had been produced in the production experiment and had to perform a rating task: Participants heard a sentence, while the two members of the minimal pair were presented on the screen on the ends of a 5-point-scale, for instance:

pond 1--2--3--4--5 pont

Participants had to type a '1' on the keyboard in front of them when they were sure they heard the word standing left on the screen (*pont* in the above example), a '2' when they thought they rather heard that word than the other one. They typed a '5' when they thought they heard the word presented on the right side of the screen (here *pond*), a '4' when they saw a trend towards this word. They typed a '3' when they could not make a decision between the words. When they wanted to listen to the sentence again, they could use a push button device. Every sentence could be repeated maximally twice. After the second repetition, the push button device did not react anymore and participants had to type their score which was automatically written to a result file. A single trial looked as follows: 50 ms after the onset of the visual information the sentence was played. The scale remained on the monitor until the participant typed in a rating response. After a pause of 1000 ms the following trial started.

Participants saw the word ending in the voiced consonant on the left side of the screen in 50% of the trials, and in 50% on the right side. In 50% of the trials the correct answer appeared on the left and in 50% of the trials on the right. However, one participant could not judge each item twice, since this would have prolonged the experiment too much. Therefore, participants belonged to two different groups. Item pairs of which the voiced member occurred on the left side of the screen for the participants of the first group had their voiceless member appearing on the right for the participants of the other group. For example, having heard the sentence *Pien zegt "Ik zie een wand en een ever"* spoken by speaker 1, participants of the first group saw the scale *wand 1--2--3--4--5 want* on the monitor, and participants of the other group saw *want 1--2--3--4--5 wand.* Speakers and items were distributed in a controlled way, see Appendix A8.3.

Participants were informed that in 50% of the cases they would hear the member of a pair ending in a 'd', and in 50% of the cases the member ending in a 't'. This was the proportion of underlyingly voiced and voiceless targets as produced in the production experiment. Participants were told that because of the procedure the sentences were made it could happen that they sounded a bit odd, and that furthermore they should concentrate on the signal, since the meaning of the sentences would not give them any cue about which member of a minimal pair was in the sentence.

After they had indicated that they understood the instruction, *Attentie* occurred on the screen for 3500 ms, followed by the twelve practice trials. Those appeared in the same randomization for all participants. Then the experimental trials were presented, which had been randomized as follows: The whole list of all items spoken by all speakers (=168 trials) was randomized once (= randomization 1). Then it was split into three parts of 56 trials each. Each of the three parts was randomized individually and concatenated, starting with the second part followed by the third and ending with the first (= randomization 2). Then the three parts were again randomized individually and concatenated in the order third - first - second part (= randomization 3). Then the whole list was randomized again, split into three parts, etc., until ten different randomizations were created, which were assigned to each participant group. After the last trial, the experiment ended with the word *Einde* "The end" appearing on the screen.

Data analysis. The responses were transformed such that a '1'-score meant a /t/ response, '2' meant a /t/?-response, '3' meant an undecided response, '4' stood for /d/?, and a '5'-score meant a /d/-response.

The picture provided by the proportion of correct responses in the different experimental conditions turned out to be very clear already in a qualitative analysis. Nevertheless, in addition the arcsine-transformed proportions[14] were submitted to a separate Anova for every response key. Since the materials contained all items of Dutch that could be tested, only analyses over subjects were run. They comprised the within-subjects variables, 'underlying voicing' (two levels: voiced and voiceless), 'context' (three levels: clitic, final, and nasal), and 'speaker' (four levels: speaker 1, 2, 3, and 4).

Results

Table 17 shows the number of responses and the proportions given in the five response categories for the three context conditions.

The proportion of undecided responses was low (11%) and was the same for all contexts (effect of context: $F(2;38) = 2.63$, $MSE = .08$, n.s.). The final context condition got less undecided responses (93 cases, these are 9% of the responses in the final context) than the other contexts (clitic context: 126 cases, corresponding to 12%; nasal context: 128 cases, corresponding to 12% of the responses in the respective context condition). But a pair comparison for repeated measurements showed that final and nasal context did not differ significantly.[15]

As expected, participants did not exceed chance level in deciding on the target word's final obstruent in the final and the nasal context. (44% correct responses for both contexts). The voice contrast is neutralized, and participants, who had been told that half of the targets were underlyingly voiced, could assign the responses only by guessing. Crucially, however, in the clitic context the proportions of correct responses also turned out to be low (42%).

In the final and clitic context, participants were more likely to give correct responses to voiceless than to voiced targets (clitic: 50% correct for voiceless, 33%

---

[14] For every subject, the frequencies per cell (context by voicing by speaker) were divided by 7 (=number of items). Proportions of 0 were replaced by 1/2n=1/14, proportions of 1 were replaced by 1-1/14. The frequencies were arsine-transformed.

[15] While some authors (e.g., Stevens, 1992[2]) use a pooled error term for all comparisons within a rating, I used the Tukey HDS procedure following Hays (see Hays, 1989 [4] , p. 529f), who calculates a different value for the error term for every comparison.

for voiced, final: 56% correct for voiceless, 31% correct for voiced), while in the nasal context, there were more correct responses to voiced than to voiceless targets (61% for voiced, 26% for voiceless), see Figure 18. This pattern was due to the higher number of voiced responses given in the nasal context: Of the 1459 voiced

	response						
condition	(1) t	(2) t?	(1)+(2)	(3) ?	(4) d?	(5) d	(4)+(5)
clitic							
voiced	165 (.29)	154 (.28)	**319 (.57)**	59 (.11)	94 (.17)	88 (.16)	**182 (.33)**
voiceless	160 (.29)	121 (.22)	**281 (.50)**	67 (.12)	106 (.19)	106 (.19)	**212 (.38)**
final							
voiced	211 (.38)	128 (.23)	**339 (.61)**	45 (.08)	91 (.16)	85 (.15)	**176 (.31)**
voiceless	200 (.36)	112 (.20)	**312 (.56)**	48 (.09)	120 (.21)	80 (.14)	**200 (.36)**
nasal							
voiced	79 (.14)	76 (.14)	**155 (.28)**	61 (.11)	176 (.31)	168 (.30)	**344 (.61)**
voiceless	70 (.13)	78 (.14)	**148 (.26)**	67 (.12)	175 (.31)	170 (.30)	**345 (.62)**

Table 17. Frequency and Proportion (in Parentheses) of the Different Response Categories in the Three Context and Two Voicing Conditions in Experiment 8a

responses given, 47% were assigned to the nasal context, 27% to the clitic, and 26% to the final context. This shows that - as expected - the nasal caused regressive voice assimilation on the preceding stop. However, the higher proportions correct responses for voiceless targets in clitic and final context were accompanied by a high proportion of wrong voiceless responses to voiced targets (57% in clitic, 61% in final context). And in the nasal context, not only voiced, but also voiceless targets got many voiced responses (62%). In sum, participants could not differentiate between voiced and voiceless targets in either context condition.

Although the qualitative analysis above already showed clear results, a quantitative analysis was conducted. Since the clear /t/- and /d/-responses and the responses in the 'likely /t/' and 'likely /d/'-categories patterned similarly in all voicing and context conditions, clear and likely responses were collapsed. In the following, the notion 'voiceless response' refers to the sum of 'clear' and 'likely' /t/-responses, and the notion 'voiced response' to 'clear' and 'likely' /d/-responses.

Since voiced and voiceless and the small proportion of undecided responses add to 100%, and the undecided responses did not differ between the two context

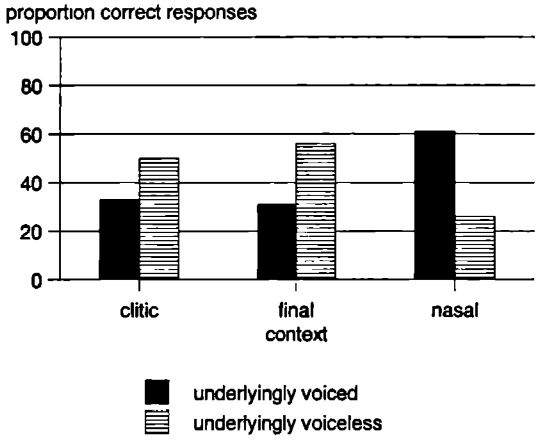


Figure 18. Proportion Correct Responses Underlyingly Voiced and Voiceless Stops in the Three Context Conditions

conditions, only the results for the voiceless responses will be reported in the statistical analysis. The results for the voiced responses can be found in Appendix 8.3.

For both underlyingly voiced and voiceless target obstruents, the final and the clitic context got a high number of voiceless responses (651 cases, 59% of the responses in the final context, 600 cases, 54% of the responses in the clitic context condition), while participants gave only 303 voiceless responses (= 27% of the responses) in the nasal context condition. The effect of context was significant ($F(2;38) = 49.91, MSE = .41, p < .01$).

A Tukey pair comparison for repeated measurements revealed that final and clitic context did not differ, while final and nasal as well as vowel and nasal context differed significantly ($p < .01$). The nasal context showed more voiced responses, as mentioned above. Importantly, clitic and final context did not differ.

The variable 'speaker' had no effect. The regional differences and the speaker's sex did not influence the ratings on the voicing of the final obstruents.

The variable 'voicing' had an effect ($F(1,19) = 8.1$, $MSE = .12$, $p < .01$). Of the 1554 voiceless responses, 813 (= 52%) were assigned to voiced targets and 741 (= 48%) to voiceless targets. The slight bias towards the wrong answer was also found for the whole set of responses: Of 3013 responses (undecided responses excluded), 1443 (= 47.89%) were correct and 1570 (= 52.11%) incorrect. However, there was no significant interaction between the variables voicing and context. Whatever the reason for the slightly higher number of incorrect responses was, they were obtained in all context conditions.

Discussion

In this experiment I investigated Final Devoicing in several contexts. Minimal pairs that only differed in the voicing of their final stop consonant, for instance, *raad* ("commission") and *raat* ("honeycomb"), had been spoken in three context conditions by participants of Experiment 8a. The members of the minimal pair could occur utterance-finally, which was called the final context (e.g., *een oom en een raad/t*), or followed by a preposition that started with a nasal (the nasal context, e.g., *raad/t/ naast een oom*), or followed by a schwa-initial enclitic pronoun (the clitic context, e.g., *raad/t en een oom*). These utterances were presented to the participants of Experiment 8b, who rated on a 5-point-scale whether they had heard the voiced (e.g., *raad*) or the voiceless (e.g., *raat*) variant.

In the final context, the stop consonant was expected to be voiceless in all words, because it was in syllable-final position, where the voice contrast should be neutralized. This is indeed what was found: Participants could not distinguish underlyingly voiced words like *raad* from underlyingly voiceless one like *raat.* A similar result had been expected for the stops in the nasal condition. Since the stop cannot form an onset with the following nasal consonant, it has to be in coda position and the voice contrast should neutralize, as in the final context condition. Again, participants behaved as had been expected. They could not distinguish the members of a minimal pair. Although they chose for the word ending in a voiced stop more frequently as compared to the final context, they did so for both, the underlyingly voiced and voiceless members of a minimal pair. Probably participants chose the voiced stop more often because the consonant received some amount of voicing from the following nasal.

Finally, there had been two different predictions for the clitic context condition, where the stop was followed by a schwa-initial function word, for example, *raat en* ("honeycomb and"). Due to encliticization, the stop will surface in the onset of the second syllable: $(\text{ra:})_\sigma(\text{tən})_\sigma$. Following an account that includes resyllabification, there will first be an abstract syllable structure for which the individual words are syllabified. Final Devoicing can apply on these syllables, for instance on the /d/ in *raad en*, before the stop resyllabifies to the onset of the following syllable. Therefore, participants should not be able to hear the difference between underlyingly voiced and voiceless stops in the clitic context. Following Levelt's model of phonological encoding, on the other hand, speakers immediately produce those syllables that surface in encliticized forms, without creating any abstract intermediate forms. Since the stop surfaces as an onset, it never is in a coda position, where it could devoice. Consequently, participants should clearly hear the difference between words ending in an underlyingly voiced and a voiceless stop. The results showed that, as in the final and the nasal context, participants could not distinguish the two members of a minimal pair in the clitic context. Final devoicing hence applies in encliticized forms.

encliticized	not encliticized	
ω	ω	ω
/ \	|	|
(ra:)$_\sigma$(**dən**)$_\sigma$	(ra:**t**)$_\sigma$	(ɛn)$_\sigma$

Figure 19. Predictions of the Model of Phonological Encoding for Final Devoicing in Encliticized and not Encliticized Forms

However, there remains a potential problem with the study described above: Encliticization is an optional process in Dutch. When it does not apply, the noun and the function word form separate phonological words. In this case, the strong form of the function word is realized and is a prosodic word on its own. Under these circumstances, also the model of phonological encoding predicts that the word-final obstruent is voiceless, since the two prosodic words are syllabified separately. The final obstruent of the noun is then realized as voiceless, as shown in Figure 19.

The very fast tempo in the production task makes it rather unlikely that many forms were not encliticized. In any case, the speech signal used in the perception task had to be checked to exclude items that were not encliticized.

One might consider acoustic measurements on the materials of the experiment to make sure that there really was encliticization in sequences like ....*raad en*. One might investigate whether the function word contained a full or a reduced vowel. Only the latter would indicate that the function word had been encliticized. Relevant cues for vowel quality might be the first and second formant frequencies and the duration of the vowel of the function word. The frequency of the first and second formant should be slightly higher for the full vowel (which should be a front, half-open [ɛ]) than for the central schwa (Kent & Read, 1992). Furthermore, the reduced vowel should be shorter. Schwa is significantly shorter than other short and long vowels in read as well as in spontaneous Dutch (Koopmans-van Beinum, 1994). As a baseline for measurements one needs as a reference point a full vowel produced by the same speaker against which the other vowels are evaluated, e.g., a vowel produced in isolation. It has been argued that the schwa-vowel in Dutch can be seen as a 'bench-mark' for a speaker from a phonetic point of view (Koopmans-van Beinum, 1982).

But whereas this seems to be suitable for investigating isolated vowels, or vowels that are surrounded by the same consonants in a controlled way, this procedure is not useful for connected speech, where we find - even for full vowels - a high amount of target undershoot, i.e. the frequencies of the vowel do not reach the values of the vowel spoken in isolation. It seems tricky to look at the acoustic properties of the vowel to investigate whether it is a schwa or not, when 'full' vowels are not always realized 'full' in connected speech. Furthermore, Koopmans-van Beinum stated that in Dutch the schwa in spontaneous as well as in read speech "seems to be the most unstable vowel of all, occupying almost the whole vowel plane" (Koopmans-van Beinum, 1994, p. 73). As an additional complication, the duration of a vowel is also affected by the surrounding consonants. Since the two vowels - /ə/ and /ɛ/ - are very close to each other in the vowel system anyhow, and given the facts described so far, phonetic measurements could not be used to distinguish between cliticized forms and full forms of function words. Instead, two phonetically trained judges were asked to investigate the materials.

Judging the Materials

Two phonetically trained judges listened to part of the materials of the perception experiment. They checked the items from the clitic context condition, in which the member of the minimal pair was followed by the schwa-initial function word (e.g., *Pien zegt "Ik zie een <u>raad en</u> een oom"*) and had to decide whether the function word was cliticized to the preceding noun. They were paid HFL 20,-.

Procedure. The sentences were presented with the same setup as in the preceding experiment. The judges could repeat a sentence by pressing a button on a pushbutton panel. They were allowed to listen to a sentence as often as they liked. After they had made a decision, they had to type a '1' on a keyboard when they thought that there was a prosodic boundary between the two words, and a '5' when they thought that the words belonged to one prosodic unit. They were told that cues for the existence of a boundary might be a pause or a glottal stop between the noun and the function word, or a full vowel of the function word. From the occurrence of a glottal stop one can conclude that the function word is not encliticized. Since the glottal stop is not necessarily produced, especially not in connected speech, its absence does not necessarily mean that there is encliticization. Kohler (1994) studied the glottal stop in German, which can occur in initial position of stem morphemes or words that start with a vowel. In wordinitial position he found most often glottalization in the subsequent vowel ('creak'), followed by a long glottal closure and complete glottal stop deletion. A glottal stop occurred in only 15% of the cases, and more often after silence than in context, and more often before a stressed than an unstressed vowel. Jongenburger and van Heuven (1991) investigated a corpus of spoken Dutch and obtained similar results. Of the 137 tokens where a word-initial vowel was preceded by a voiceless stop, only 33% included a glottal stop. The proportion was slightly higher after voiced sonorants (121 tokens, 45% glottal stops). As in Kohler's study, glottal stops were more likely before stressed vowels (70% glottal stops) than before unstressed vowels (32% glottal stops). In the materials of my experiment, the glottal stop would be before an unstressed vowel. Following the studies just mentioned, it cannot be regularly expected in the cases where there is no encliticization.

Judges had no other information about the sentences than the auditory signal. After they typed in their answer, the following sentence was played on the headphones.

	speaker			
	1	2	3	4
voiced	8 (.57)	1 (.07)	6 (.43)	8 (.57)
voiceless	12 (.86)	0 (.00)	7 (.50)	7 (.50)

Table 19. Frequencies and Proportions (in Parentheses) of Sentences that were Judged as Encliticized for the four Speakers and Voiced and Voiceless Targets

The judges heard one speaker after the other. The order of presentation was random within speakers. Two members of a minimal pair did not follow immediately on each other. The presentation of every speaker started with four practice sentences which were filler sentences from the production experiment. This was to enable the judges to get used to the speaker's voices. After the materials of all speakers had been judged once, they were presented for a second time. The sequence of speakers was the same, but the randomization within speakers was changed. The four practice trials at the beginning of each speaker remained the same. The responses for the second presentation were analyzed.

Data analysis. Due to the small number of judges, no statistics was run on the judgement results. The judgements should give qualitative information about the nature of the materials.

Results. The results showed that the speakers seemed to have produced encliticized forms about half of the time, except for speaker 2, who got almost no 'clitic' judgements (see Table 19).

	response						
condition	(1) t	(2) t?	(1)+(2)	(3) ?	(4) d?	(5) d	(4)+(5)
clitic							
voiced	130 (.31)	126 (.30)	**256** **(.61)**	40 (.10)	68 (.16)	56 (.13)	**124** **(.30)**
voiceless	120 (.29)	98 (.23)	**218** **(.52)**	43 (.09)	79 (.17)	80 (.17)	**159** **(.34)**
final							
voiced	155 (.33)	99 (.21)	**254** **(.55)**	36 (.08)	66 (.14)	64 (.14)	**130** **(.28)**
voiceless	141 (.30)	93 (.20)	**234** **(.51)**	40 (.09)	87 (.19)	59 (.13)	**146** **(.32)**
nasal							
voiced	54 (.12)	49 (.11)	**103** **(.22)**	46 (.11)	141 (.30)	130 (.28)	**271** **(.59)**
voiceless	47 (.10)	54 (.12)	**101** **(.22)**	50 (.11)	142 (.31)	126 (.27)	**268** **(.58)**

Table 20. Frequency and Proportion (in Parentheses) of the Different Response Categories in the Three Context and Two Voicing Conditions in Experiment 8b, Speaker 2 Excluded

Reanalysis of the Perception Experiment. Speaker 2, who was judged as producing almost no encliticized forms, was excluded from the analyses of the results of the perception experiment. Table 20 shows the results of the perception experiment without speaker 2. Notice, however, that Table 20 still includes all non-cliticized renderings of the other speakers, which were not excluded to provide more power for the quantitative analysis. This can be justified: In a conservative first analysis, only those items had been considered which both judges had judged as encliticized in the clitic context both in the underlyingly voiced and voiceless condition. These were four minimal pairs spoken by speaker 1, two pairs produced by speaker 3, and three pairs spoken by speaker 4. The distribution of responses in the three context and two voicing conditions (given in Appendix A8.4) closely resembled the pattern obtained when non-cliticized renderings were included. Therefore, the quantitative analysis was run over the latter data set.

In this analysis, the proportion of undecided responses again turned out to be low and did not differ between the three contexts ($F < 0$). Of the 840 responses that were collected in each context, the final context again got somewhat less undecided responses than the other contexts (clitic: 83 cases, 10% of the responses in this context, final: 76 cases, 9% of the responses, nasal: 96 cases, 11% of the responses). A Tukey test of paired comparisons showed that this difference was not significant.

Again, the proportion of correct responses was low in all contexts (clitic: 41%, final: 43%, nasal: 44%). As in the previous analysis, participants assigned more voiced responses in the nasal context condition than in the other contexts, for both, voiced and voiceless targets.

Since the single 'clear /t/'- and 'likely /t/'-responses again patterned similarly, as did the voiced single response categories, the quantitative analysis was run over the combined responses. The arcsine-transformed proportions were submitted to an Anova with three within-subject variables 'context' (three levels: clitic, final, and nasal), 'voicing' (two levels: voiced and voiceless), 'speaker' (three levels: speaker 1,3, and 4). As in the analyses with all four speakers, context had an significant effect ($F(2;38) = 54.62$, $MSE = .36$, $p < .01$). Pair comparisons with the Tukey test showed that final and nasal ($p < .05$) as well as clitic and nasal context ($p < .01$) differed significantly from each other, due to the high number of voiced responses in the nasal context. However, clitic and final context again did not differ significantly. Also, the variable 'voicing' again showed a significant effect ($F(1;19) = 8.9$, $MSE = .11$, $p < .01$). Of the 1166 voiceless responses, 613 (= 53%) were assigned to voiced, and 553 (= 47%) to voiceless targets. As in the analysis with four speakers, this slight bias towards the wrong response was also found in the whole set of responses: Of 2265 responses (undecided responses excluded), 1078 (= 47.59%) were correct and 1187 (= 52.41%) incorrect. Again, the variables 'context' and 'voicing' did not interact.

One might argue that the similar pattern for clitic and final context might be caused by the fact that the three speakers were judged as using encliticized forms in 57% of the items. However, 52% of the voiced targets in the clitic context were encliticized. According to the model of phonological encoding, for these the proportion of correct responses should be high, since an underlyingly voiced stop should remain voiced in encliticized forms. Conversely, participants should not exceed chance level for the 48% of the underlyingly voiced targets that were not encliticized, since the voice contrast should be neutralized, because the stop is in coda position. In total, the proportion of correct responses in the clitic context condition should then be about 75%. This however was not found.

In sum, when the speaker who used no clitics at all was excluded, clitic and final context again patterned the same. This result had been expected following an account according to which Final Devoicing applies on an intermediate abstract level of syllable structure which is later resyllabified to result in the surface syllables.

Experiment 9a: The Second Production Experiment

Although the reanalysis showed the same effects as the earlier one, there are some arguments for the need to replicate the findings of the first study.

Firstly, participants in the perception task had been told that there were 50% voiced and 50% voiceless targets. In the nasal context (e.g., *...een pond naast een oom*), where the obstruent was likely to be voiced due to regressive assimilation processes from the following nasal, participants could easily get rid of their '+voice'-responses by assigning them to this condition. It is possible that if this context had not been available, they would have listened more carefully to the remaining two contexts and have been able to pick up some differences. To exclude the possibility that participants assign their responses based on the context rather than the acoustic signal, the presentation of the different contexts should be blocked in the new perception study.

Secondly, it remains to be shown that the experiment is sensitive enough to pick up any voicing differences at all. Therefore, a plural condition should be included. In this context, participants should be able to determine the voice quality of a stop, since the voice contrast is not neutralized (e.g., *pon*[d]*en* - *pon*[t]*en*).

The third problem is the optionality of cliticized forms: It seems to be rather difficult to elicit cliticized forms consistently in a delayed naming experiment, even when participants are under a considerably high time pressure. One should choose a different task. Because of the difficulty in eliciting cliticized forms consistently in the production experiment, one could ask a speaker to produce the materials using cliticized forms consistently. However, it seems problematic to instruct a speaker how to produce utterances without biasing in the hypothesized direction. The speaker is also likely to pay more attention to the materials than the participants in the production experiment, who unknowingly produced the materials for the perception experiment. Therefore, the materials should again be produced by participants in a production experiment.

A second set of experiments was run which used a repeated articulation task. A plural context condition replaced the nasal context, and more minimal pairs could be included in the perception experiment.

Method

Participants. Twenty students, 10 men and 10 women, participated in the experiment.

Stimuli. Again, minimal pairs were tested in three context conditions. The carrier sentences differed from the first study. Instead of long sentences, the stimuli consisted of the minimal pair target noun, the conjunction *en* ("and") and a second noun, which was - like the minimal pair noun - a monosyllabic word. An example is given in Table 21.

	underlying voicing	
context	voiced	voiceless
clitic	*pond en olm*	*pont en olm*
final	*olm en pond*	*olm en pont*
plural	*ponden en olmen*	*ponten en olmen*

Table 21. Carrier Sentences in Experiments 9a,b, Examplified for the Pair *pond - pont* ("pound - ferry")

The second noun started with a vowel or with the glottal fricative /h/. A thirteenth minimal pair was included, which had not been in the set of materials in the first study. The structure of the new target sequences, with no article or inflected complementizer preceding the noun, allowed for all thirteen minimal pairs to be included, also mass nouns. In addition, plural forms were produced for the four minimal pairs of which both members have a regular plural.

Twenty-four filler items consisted of pairs of phonologically similar words (rhyming, alliterating). Ten of them contained a word ending in an underlying /d/ (but not having a minimal /t/-pair). These ten and three other fillers also occurred in the plural condition. The remaining eight fillers only occurred in clitic and final condition. As the experimental targets, the fillers were combined with monosyllabic second nouns that also resembled each other. For the materials see Appendix A9.1.

Apparatus. The apparatus was the same as in the first production experiment. Of the tones of different frequency that were written to one channel of the DAT-tape, one indicated the onset of the cue presentation, and the other a warning beep that was presented to the participants.

Design. The design included three context conditions. All items occurred in a final and a clitic context, and four items in addition in a plural context. Moreover, items could be either underlyingly voiced or voiceless.

Procedure. Participants were told to memorize the sequence that was presented on the screen and to reproduce it as quickly as possible when they saw a cue signal. They were told that they had to complete their response before they heard a warning tone over headphones and that the cue would reappear and they had to repeatedly reproduce the target sequence.

A single trial looked as follows: The target sequence appeared in the center of the screen for 1000 ms. After a pause of 250 ms, a sequence of three stars '***' appeared on the screen as a cue. After 1200 ms (measured from cue onset) participants heard a beep of 150 ms length as a signal at which their productions should be finished. The cue disappeared with beep onset. After another pause of 250 ms length, the cue appeared again and participants had to repeat the same target sequence. The time between cue onset and beep was 70 ms shorter than before. This was repeated with the time lag between cue onset and beep decreasing stepwise by 70 ms until it was 430 ms short. This resulted in 11 productions of every target. After a pause of 250 ms, the following trial began.

There were 60 experimental targets (13 pairs x 2 voicing x 2 context = 52, + 4 pairs x 2 voicing in plural condition) and 64 fillers (16 fillers x 3 context, + 8 fillers x 2 context). Six filler items (two of each context condition) occurred as practice trials after the word *Attentie* ("attention") had appeared on the screen for 3500 ms. After the six practice trials, the remaining 112 trials were presented in randomized order. The items were randomized following the procedure used in the first production exepriment. At least three trials separated two members of a minimal pair. Maximally three successive trials appeared in the same context condition. The same filler item did not occur without at least one intervening trial. Then the word *Pauze* ("pause") occurred on the screen. During the experiment, the experimenter had marked the targets where an error had occurred (e.g., coughs, too hard or quiet voice, stuttering). Items with an error were repeated at the end of the experiment.

The duration of the experiment was about 30 minutes. As in the first perception experiment, participants had to fill in a question form after the experiment.

Experiment 9b: The Second Perception Experiment

In the second perception experiment, participants again rated the voice quality of the noun's final consonant on a 5-point-scale, and the procedure was the same as in the first perception task. As in the first perception experiment, they were told that the materials contained 50% /d/-targets.

Method

Participants. Sixteen students (8 men, 8 women) participated.

Stimuli. Participants listened to thirteen minimal pairs that occurred in a clitic and a final context, and four minimal pairs in a plural context, spoken again by two men and two women, two from the South and two from the rest of the Netherlands (= 240 experimental trials). Of the 11 repetitions that had been produced in the production experiment, normally the fifth was chosen for the perception task. Only in cases where this repetition contained an error, the sixth or fourth repetition was taken. Again, the utterances were manipulated by replacing the second noun. The noun *olm* ("elm") appeared as a second noun in all sentences.

Apparatus. The apparatus was the same as in the first perception experiment.

Design. The design included four within-subject variables: 'Underlying voicing' (two levels: voiced and voiceless), 'speaker' (four levels for the four speakers), 'context' (two levels final and clitic for all minimal pairs, and an additional level plural for four of the pairs).

In contrast to the former study, the presentation of contexts was blocked. One group of participants started with the clitic condition, followed by the final condition, the other group started with the final condition, followed by the clitic condition. Both groups ended with the plural condition. The blocked presentation should prevent participants from assigning the responses on the basis of the context variation. Within the two order groups, there were again two groups of participants: The items in each context condition were assigned to two blocks of 52 productions. In every block, the two members of the minimal pair occurred twice, spoken by different speakers. Once the voiced member of the pair appeared on the left side of the screen, once on the right side. Consequently, there were 50% of the trials with the correct response on the left, and 50% with the correct response on the right side of the screen. There were 13 productions for each of the four speakers in one block. Half of the participants starting with the clitic condition started with block 1 followed by block 2. For the other participants, the order of blocks within a context condition was reversed.

As in the first perception experiment, the right-left assignment of the members of a minimal pair on the scale was counterbalanced for two participant groups (see Appendix A9.2).[16] The same voicing condition followed each other maximally five times, and not more than three times the same speaker occurred in consecutive trials.

Procedure. The structure of a single trial was the same as in the First Perception Experiment. The experiment started with a practice block which was the same for all versions. It contained pairs of phonologically similar words like *hals - halm*, two in the final context condition, three in clitic, one in plural. Those six items were produced twice. In total there were three productions of every speaker in the practice block. Four pauses interrupted the experiment, two between the three context conditions, and two at the end of the first block in the clitic and the final context condition. Participants determined the length of the pauses. The perception experiment took about 35 minutes.

Data analysis. The same analyses were conducted as in the First Perception Experiment. The level plural replaced the level nasal of the variable 'context'.

Results and Discussion

Table 22 shows the number of responses and proportions. As in the first perception experiment, the proportion of undecided responses was low. However, this time there was an effect of context for the undecided responses ($F(2;30) = 7.11$, $MSE = .12$, $p < .01$). This was due to the plural context condition, where participants gave almost no undecided responses. The proportion of undecided responses in the clitic and final context was on average 9% and did not differ. The interaction of context and voicing was not significant.

As in the first perception experiment, first the amount of correct responses will be described. Then follows the qualitative analysis of the voiceless responses.

The accuracy was high in the plural context condition, for both voiced and

[16]In total, there were then sixteen different orders: Participants starting with the clitic condition and with the block order 1-2, participants starting with the clitic condition and with block order 2-1, participants starting with the final condition and block order 1-2, participants starting with the final condition and the block order 2-1. Of those four different orders there were two versions each, where randomizations within the blocks were different in every stackfile. Those eight different files were copied and the speakers were reversed for the second group of participants.

voiceless targets, which yielded 95% and 91% correct responses, respectively. The other context conditions had low proportions of accurate responses, 49% for the final and 45% for the clitic context. In both clitic and final context, participants were more likely to give correct responses to voiceless than to voiced targets (final context: 56% correct for voiceless, 41% for voiced; clitic context: 50% correct for voiceless, 40% for voiced). This is shown in Figure 20. The detection rate for voiced targets was similar in both contexts and slightly below chance (41 vs. 40%).

	response						
condition	(1) t	(2) t?	(1)+(2)	(3) ?	(4) d?	(5) d	(4)+(5)
clitic							
voiced	273 (.33)	151 (.18)	**424** **(.51)**	79 (.09)	138 (.17)	191 (.23)	**329** **(.40)**
voiceless	272 (.33)	147 (.18)	**419** **(.50)**	87 (.10)	131 (.16)	195 (.23)	**326** **(.39)**
final							
voiced	273 (.33)	148 (.18)	**421** **(.51)**	67 (.08)	142 (.17)	202 (.24)	**344** **(.41)**
voiceless	302 (.36)	162 (.19)	**464** **(.56)**	56 (.07)	123 (.15)	189 (.23)	**312** **(.38)**
plural							
voiced	9 (.04)	5 (.02)	**14** **(.05)**	0 (.00)	14 (.05)	228 (.89)	**242** **(.95)**
voiceless	212 (.83)	21 (.08)	**233** **(.91)**	7 (.03)	7 (.03)	9 (.04)	**16** **(.06)**

Table 22. Frequency and Proportion (in Parentheses) of the Different Response Categories in the Three Context and Two Voicing Conditions in Experiment 9b

Already the qualitative description showed that - given the low detection rate for voiced targets in the clitic context - voiced stops were apparently devoiced in the clitic context. As in the final context, participants could not detect them. Clitic and final context hence again did not differ, whereas participants could distinguish underlyingly voiced and voiceless stops in the plural context. In the plural context, participants gave almost only extreme responses (keys 1 and 5), indicating that they were more confident in their decisions than in the other contexts.

91% of the voiceless targets in the plural context got voiceless responses, and 95% of the voiced targets got voiced responses. In the final context, 51% of the voiced targets got voiceless, 41% voiced responses. For the voiceless targets there were 56% voiceless and 38% voiced responses. Finally, in the clitic context, 51% of the voiced and 56% of the voiceless targets were responded to as voiceless. Voiced responses were given to 40% of the voiced and 39% of the voiceless targets.

There was no effect of context. Participants assigned about 50% voiceless responses in all contexts (clitic: 843 cases, 51% of the responses in this context condition, final: 885 cases (53%), plural: 247 cases (48%)).

The quantitative analysis showed that the effect of underlying voicing was significant ($F(1;15) = 275.12$, $MSE = .02$, $p < .01$). The proportion of voiceless responses was higher for voiceless targets (1116 cases, 57% of the voiceless responses) than for voiced targets (859 cases, (43%)). This effect was caused by the high amount of correct responses in the plural condition. The interaction of voicing and context was significant, too: ($F(2;30) = 281.12$, $MSE = .03$, $p < .01$). Tests of simple effects revealed that the interaction was due to a significant difference in the plural condition, where voiced targets yielded almost only voiced, and voiceless targets almost only voiceless responses.

The pattern in the clitic context condition did not differ from the one in the final context condition. In both contexts, participants assigned voiceless responses to about half of both, the voiced and the voiceless targets, indicating that they could not distinguish between the two members of a minimal pair. Final devoicing must

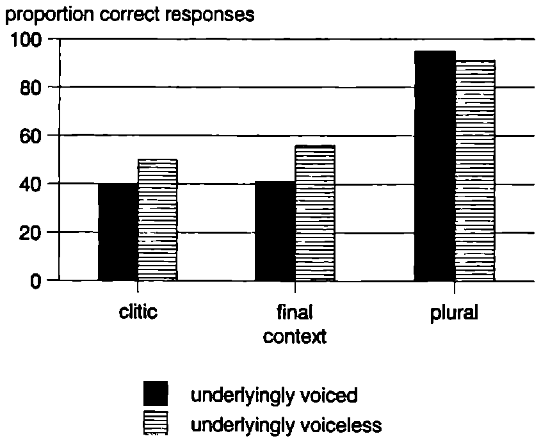


Figure 20. Proportion Correct Responses for Underlyingly Voiced and Voiceless Targets in the Three Context Conditions have taken place in the final as well as in the clitic context. In the plural context, however, participants were able to correctly identify voiced and voiceless targets. This was expected since there is no devoicing in the plural forms. The results in the plural condition furthermore showed that participants were in principle capable to distinguish the voicing contrast in this experiment.

The variable 'speaker' had a significant effect ($F(3;45) = 26.17$, $MSE = .04$, $p < .01$). Generally, in the clitic and final conition, female voices got less voiced responses than male voices. This is in harmoy with Slis' investigations on fricatives in Dutch, where he found less voicing in female voices. Furthermore, while the targets produced by speakers 1 (female, from the South), 2 (male, from the North), and 3 (female from the North) yielded more voiceless than voiced responses, the opposite was the case for speaker 4, a male speaker from the South of the Netherlands. This is shown in Table 23.

	response						
speaker	(1) t	(2) t?	(1)+(2)	(3) ?	(4) d?	(5) d	(4)+(5)
1	412 (.43)	187 (.19)	**599** **(.62)**	60 (.06)	95 (.10)	196 (.20)	**301** **(.31)**
2	331 (.34)	133 (.14)	**464** **(.48)**	77 (.08)	144 (.15)	275 (.29)	**419** **(.44)**
3	350 (.36)	195 (.20)	**545** **(.57)**	68 (.07)	119 (.12)	228 (.24)	**347** **(.36)**
4	248 (.26)	119 (.12)	**367** **(.38)**	91 (.09)	187 (.19)	315 (.33)	**502** **(.52)**

Table 23. Frequency and Proportion (in Parentheses) of the Different Response Categories For the Four Speakers in Experiment 9b

The variables 'speaker' and 'context' interacted significantly ($F(6;90) = 11.35$, $MSE = .04$, $p < .01$). Table 24 shows the results in the combined voiced and voiceless response categories for the four speakers and three context conditions.

Speaker 4 was perceived as producing mainly voiced stops in the clitic context condition (see the bold numbers above), while he patterned as the other speakers in the other contexts. The interaction of context and voicing was significant within every speaker, also speaker 4 (speaker 4: $F(2;30) = 58.36$, $MSE = .05$, $p < .01$). Importantly, however, the bias towards voiced responses for speaker

4 in the clitic context was equally distributed over both voiced and voiceless targets, see the bold numbers Table 25. Thus, speaker 4 produced more voicing, but he did so for both, the underlying voiced and voiceless targets.

	context		
speaker	clitic	final	plural
	response t (1)+(2)		
1	260 (.63)	273 (.66)	66 (.52)
2	202 (.49)	202 (.49)	60 (.47)
3	260 (.63)	229 (.55)	56 (.44)
4	**121 (.29)**	181 (.44)	65 (.51)
	response d (4)+(5)		
1	111 (.27)	130 (.31)	60 (.47)
2	172 (.41)	181 (.66)	66 (.52)
3	117 (.28)	160 (.38)	70 (.55)
4	**255 (.61)**	185 (.44)	62 (.48)

Table 24. Frequency and Proportion (in Parentheses) of the Combined Response Categories in the Three Context Conditions for the Four Speakers in Experiment 9b

	context		
voicing	clitic	final	plural
voiced	**129 (.62)**	89 (.43)	60 (.94)
voiceless	**126 (.61)**	96 (.46)	2 (.03)

Table 25. Frequencies and Proportions (in Parentheses) for the Combined 'Voiced' Response Category for Speaker 4 in the Three Contexts and Two Voicing Conditions

To summarize, participants could correctly identify voiced and voiceless targets in the plural context. This shows that they are in principle able to distinguish between voiced and voiceless stops. However, the detection of voiced targets is slightly below chance in both final and clitic context. Stops in encliticized forms are apparently devoiced.

Acoustic Measurements

In the perception experiments, participants were unable to distinguish underlyingly voiced and voiceless stops when these occurred in a final or a clitic context. However, this does not necessarily have to mean that the acoustic signal does not differ for underlyingly voiced and voiceless stops. Perhaps it does in fact differ, but participants were just unable to use this difference for their rating decision in the perception task. To investigate whether the acoustic signal provides any cues to the underlying voice contrast, the stimuli of the Second Perception Experiment (= Experiment 9b) were examined.

As has been discussed in the section on Experimental Studies on Final Devoicing above, there have been several phonetic studies that investigated final devoicing. They are concerned with the relation of phonology and phonetics. Phonologists suggested that the underlying voice contrast is neutralized in final position. But this leaves open the question in how far this neutralization is realized phonetically. Some researchers (e.g., Jassem & Richter, 1989; Fourakis & Iverson, 1984) found that underlyingly voiced and voiceless stops did not differ acoustially when they were in final position. This suggests that the neutralization process is complete at the phonetic level. Other studies, however, reported an acoustic difference between underlyingly voiced and voiceless stops in final position and argued that neutralization is incomplete (e.g., Slowiaczek & Dinssen, 1985; Port & O'Dell, 1985). Most of the studies use a set of the following variables that can provide cues for voicing (for a list of cues for voicing see, for instance, the summary in Kent and Read, 1992).

The first measure for voicing refers to the temporal coordination of the production of the stop consonant with the vibrating vocal folds that provide voicing for a following vowel. The production of a stop consonant involves two stages. When producing a stop, a speaker completely closes the vocal tract - the closure of the lips to produce a [p] is an obvious example. As a result, pressure arises in the mouth, which is released when the closure is released. In many cases, the acoustic manifestation of this release is a burst, that is a short fricative noise. Importantly, the release of the closure and the vibration of the vocal folds are coordinated differently for voiced and voiceless stops. For voiceless stops, the vibration of the vocal folds starts relatively late after the release of the stop closure. For voiced stops, voicing begins at an earlier point of the production of the stop consonant. Voicing starts either simultaneously or immediately after the stop closure, or even before the release - the latter option is called 'prevoicing'. The time interval between the onset of voicing and the release of the stop is called the voice onset time (VOT). It is a useful measure for stops that are preceded by silence, e.g. the contrast in syllables like [pa] and [ba]. When stops are in intervocalic position, one cannot determine the start of voicing.

A second variable that differs between voiced and voiceless stops is the length of the vowel preceding the stop consonant, which tends to be longer before voiced than before voiceless stops. A third difference is the longer closure duration of voiceless as compared to voiced stops. Fourthly, voiced stops have a shorter release burst or do not have a release burst at all. By the shorter duration of closure and release, a voiced stop consonant is shorter than a voiceless one. A fifth cue to voicing can be provided by the formant transitions of the preceding and following vowel. With respect to the length of the transitions, they develop fully in time for vowels neighboring voiced stops, while they are cut down and start late for the voiceless stops. Furthermore, the first formant frequency shift from the first vowel to the stop and from the stop to the second vowel is rather large in the case of voiced stops, and smaller with voiceless stops. A sixth cue to voicing is voice activity during closure, which is present in voiced stops. Voicing during closure manifests itself in a voice bar (= energy of the first formant). Voiceless stops lack a voice bar, they have a silent period during closure. For a detailed description of the acoustic features of voiced and voiceless stops in Dutch, see Slis (1985), Kuijpers (1993a,b).

Not all of these variables could be used for investigating the stimuli of the perception Experiment 9b. Because VOT cues voicing in initial stops, but the stops in the stimuli were in intervocalic position, it was not measured. Moreover, the utterances had been produced in a high tempo. This makes it unlikely to pick up any VOT differences reliably, because VOT values are considerably small. They range from about -20 ms to 20 ms for voiced, and from 25 ms to up to 100 ms for voiceless stops (see Kent & Read, 1991, p. 108ff). These values were obtained with carefully spoken monosyllabic tokens, for the present stimuli they are expectedly smaller. Similarly, the slope of the transitions could not be determined reliably.

The acoustic measurements included the duration of the vowel preceding the stop, the closure duration, the duration of the burst and of the whole consonant, and the duration of the transitions of the preceding and following vowel. Furthermore, I determined whether a voice bar was present or absent.

Method

Stimuli. The stimuli were the materials of Experiment 9b. They had been spoken by four participants of Experiment 9a. The stimuli included 13 minimal pairs (e.g., *raad* "commission" - *raat* "honeycomb") in each the final and the clitic context (final: *olm en raad/t*, clitic: *raad/t en olm*). In addition, four pairs occurred in a plural form (plural: *rad/ten en olmen*), resulting in 240 items (= 13 pairs x 2 voice x 2 contexts x 4 speakers, plus 4 pairs x 2 voice x 4 speakers).

Apparatus. The stimuli were stored digitally on a Silicon Graphics UNIX computer, using 20 kHz sample frequency. They were analyzed with the speech processing program XWAVES.

Procedure. Measurements were done separately for each speaker. The following measurements were done:

The length of the vowel that preceded the obstruent was measured from the first regular vocalic period either after the burst or friction noise of the preceding stop or fricative, or, in the case of preceding sonorants, when the waveform changed abruptly. The end of the vowel was at the onset of the silent period of the stop consonant. When items ended in a cluster, which was the case for the pairs *bond/t, ford/t, pond/t, wand/t*, the consonant that intervened between the vowel and the stop was included in the vowel segment.

The length of the total voiceless stop consonant was measured from the onset of the silent period. For voiced stops, their total duration was measured from changes in amplitude and formant structure in the waveform and the spectogram. The end of the consonant was determined by the end of the burst. If no burst was present, the end was marked by the regular vocalic waveform of the following vowel.

The duration of the burst reached from the onset of the friction noise until its end. The closure duration was not measured separately. Instead, it was calculated by subtracting the duration of the burst from the duration of the whole stop consonant.

The length of the formant transitions in the vowels preceding and following the obstruent were measured on the basis of the first formant frequency shift in the spectogram and simultaneously controlled by looking for changes in the waveform. The transition of the preceding vowel was only measured for the words that had no coda clusters (these were the pairs *bod/t, graad/t, kruid/t, lood/t, nood/t, raad/t, rad/t, lab/p, slib/p*). Since no vowel follows the obstruent in the final context, there is no formant transition into a following vowel.

The absence or presence of voice activity was determined from the absence or presence of low frequency energy in the spectogram and the waveform. This was again measured only for words without consonantal clusters.

An example of a labelled item is shown in Appendix A9.4.

Data analysis. I will refer to the measured units like the burst or the preceding vowel as 'segments'. The mean segment durations were calculated for every segment by context by voice by speaker cell. Missing values, which were caused by noise in the signal that prevented exact measurements, were replaced by cell means. The 15 missing values (= 1.6% of the whole data set), were equally distributed over the voiced and voiceless categories at a particular context condition (see Table 26). The duration of the stop closure was calculated by subtracting the burst duration from consonant duration for each item. Furthermore, two of the four minimal pairs that had contributed to the plural context had to be excluded from the measurements. In these two pairs, a nasal preceded the stop. Especially the voiced stops were difficult to measure in this phonetic context. In the remaining two pairs that had been produced in the plural context a vowel immediately preceded the stop consonant, so that the stop consonants could be measured reliably.

The durational values of the measured segments were submitted to four separate Anovas for preceding vowel, closure, burst, and consonant. Since the formant transitions turned out to be similar in all context and voicing conditions, they were not quantitatively analyzed. Moreover, the plural forms were not included in the Anova, because there were not enough items available in this context. The four Anovas comprised the variables 'context' (two levels: clitic and final), 'voicing' (two levels: voiced and voiceless), and 'speaker' (four levels for the four speakers), and treated the 13 pairs as subjects.

Results and Discussion

Table 26 shows the mean length of the segments and the standard deviations, the number of items that contributed to each segment, and the number of missing values. Furthermore, the number and proportion of items where a voice bar was present in the spectogram is given. The mean durations for the preceding vowel, the stop closure, and the burst in the different voicing and context conditions are depicted in Figure 21.

The durational values for the formant transitions preceding and following the stop consonant were very small. The values for the preceding formant transition did not differ for voiced and voiceless items in the clitic and final context (14 ms for voiced and 13 ms for voiceless stops in the clitic context, and respectively 13 ms and 14 ms in the final context). Formant transitions preceding voiced stops were slighly shorter than those preceding voiceless stops in the plural condition (9 ms for voiced, 12 ms for voiceless stops). This is the opposite pattern than reported in the phonetic literature (see e.g., Slis, 1986). Voiced stops should be preceded by longer transitions than short ones. But the segments were very short, and probably did not influence the perception of voicing to a large extent. This definitely holds for the 1 ms difference in the clitic and final contexts.

The values for the following formant transition did again not differ for voiced and voiceless stops in the clitic and plural context (8 ms for voiced and voiceless stops in the clitic context, 7 ms for voiced and 9 ms for voiceles stops in the plural context), and were therefore not further considered. The formant transitions

		context					
		clitic		final		plural[a]	
		underlying obstruent					
		/d/	/t/	/d/	/t/	/d/	/t/
segment							
preceding	**length[d]**	**105**	**102**	**153**	**150**	**125**	**118**
vowel	SD[d]	29	31	44	44	36	27
	n[e]	52	52	52	52	8	8
closure	**length**	**29**	**26**	**46**	**52**	**33**	**33**
	SD	16	14	33	28	6	10
	n	52	52	52	52	8	8
burst	**length**	**34**	**36**	**89**	**89**	**4**	**45**
	SD	15	13	30	33	8	17
	n	50	50	51	52	8	8
	missing	2	2	1			
stop	**length**	**63**	**61**	**135**	**140**	**37**	**78**
	SD	20	15	33	36	7	20
	n	52	52	52	52	7	8
	missing[f]					1	
preceding	**length**	**14**	**13**	**13**	**14**	**9**	**12**
formant	SD	18	6	8	6	4	5
transition[b]	n	35	35	36	35	7	8
	missing	1	1		1	1	
following	**length**	**8**	**8**	-	-	**7**	**9**
formant	SD	6	4	-	-	4	2
transition	n	47	52	-	-	7	7
	missing	5				1	1
voice bar[b]	**present**	**5**	**4**	**1**	**0**	**8**	**1**
	proportion[c]	.14	.11	.03	0.0	1.0	.12

[a]items without coda clusters (2 minimal pairs)
[b]items without coda clusters (9 minimal pairs)
[c]of 108 items in clitic and final, and 24 items in plural context
[d]in ms, after missing values replacement, [e/f]number of items/missing values in cell

Table 26. Durational Values and Presence of Voice Bars for the Stops in the Three Context and Two Voicing Conditions.
Note. The values for length and standard deviations represent ms.

preceding and following the stop consonant gave no cue for voicing, at least for the stimuli under investigation, which were produced at a rather high tempo.

The longest durations for the remaining segments (preceding vowel, stop

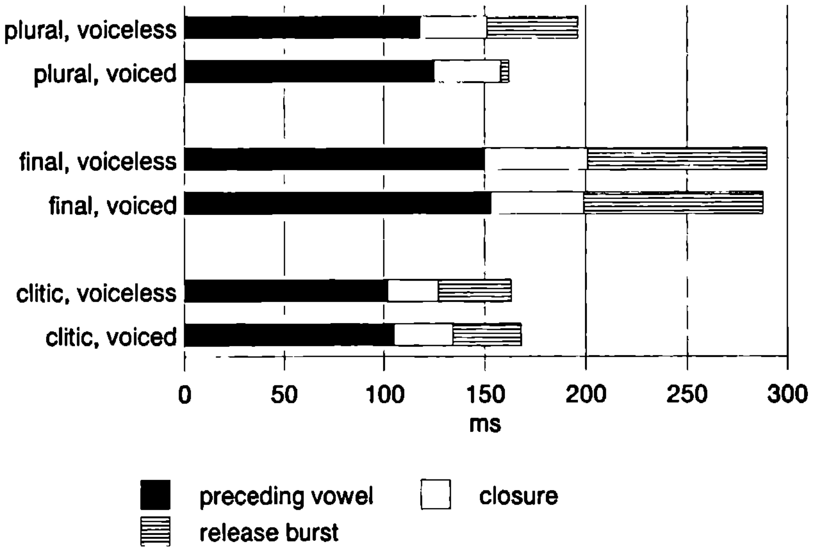


Figure 21. Duration (in ms) of Preceding Vowel, Stop Closure, and Release Burst for Underlyingly Voiced and Voiceless Stops in the Three Context Conditions

closure, burst, and the whole stop consonant) were found in the final context condition, for instance, *olm en raad/t.* This reflects the general finding of utterance-final lengthening (e.g., Lehiste, 1973; Oller, 1973; Cooper & Paccia-Cooper, 1980; Kohler, 1986). Lehiste (1973) showed, for instance, that the final metrical foot in a sentence like *Why did Linda contradict him* was longer than when it stood in other positions in the sentence.

Importantly, only the plural context showed an effect of underlying voicing. A vowel preceding a voiced stop was slightly longer than a vowel preceding a voiceless stop (125 ms and 118 ms, respectively). This trend was rather small (7 ms), given that the standard errors were 13 ms for voiced and 10 ms for voiceless plural items. But the trend was in the direction as predicted by the phonetic literature: Voiced stops should be preceded by longer vowels than voiceless ones. The whole voiceless stop consonant, however, had more than two times the duration of the voiced stop consonant. This was caused by the burst, which was almost absent in the voiced stops, but fairly pronounced in the voiceless stops. Together with the closure duration, which did not differ for voiced and voiceless stops, this adds up to much longer voiceless stop consonants. The durational patterns of voiced and voiceless stops hence differed substantially in the plural forms. These findings were supported by the presence of voice energy during closure, which is a cue for voicing, in 100% of the voiced, but only 4 % of the voiceless stops in the plural condition. Conversely, the final context showed a voice bar in only 1.8% of the underlyingly voiced and in none of the voiceless items.

The durational patterns of the final and the clitic context were quantitatively analyzed. The Anovas showed that final and clitic context differed significantly. Context affected the length of the preceding vowel ($F(1;13) = 202.24$, $MSE = 58851$, $p < .01$), the duration of the closure ($F(1;13) = 8.96$, $MSE = 44150$, $p < .02$), the burst ($F(1;13) = 195.69$, $MSE = 77453$, $p < .01$), and the whole stop consonant ($F(1;13) = 180.74$, $MSE = 165293$, $p < .01$). All segments were longer in the final than in the clitic context.

Importantly, none of the Anovas revealed an effect of underlying voicing or an interaction of context with voicing. Tests of simple effects showed that, as expected, the segments of underlyingly voiced and voiceless stops were equally long in the final context, for instance, *olm en raad/t*. An exception was the stop closure duration. The 6 ms difference between voiced and voiceless stop closures (46 ms and 51 ms, respectively) was significant on .05-level and in the direction as predicted by phonetic theory: Voiceless stops had longer closures. Piroth, Schiefer, Janker and Johne (1991) reported a similar result for German, where the closure and the release of voiceless stops turned out to be longer as compared to voiced stops in utterance-final position. The results in the perception study, however, suggest that this difference alone does not suffice to signal a voice difference.

In a clitic context like *raad/t en olm*, the underlying voice contrast was completely neutralized. This finding again resembles the results of the study by Piroth et al., who found a complete neutralization of the voice contrast for morpheme-final stops. When a stop preceded a morpheme boundary, the voice contrast was completely neutralized in terms of the duration of preceding vowel, closure, and release. The voice quality of the phoneme that followed the stop consonant did not influence the findings. Neutralization of the voice contrast was found, for instance, in *ratlos* ("helpless") and *Radlager* ("wheel bearing"), where the stops are followed by a (voiced) labial, as well as in *Ratschlag* ("piece of advice") and *Radfahrer* ("cyclist"), where the stops are followed by a voiceless fricative.

A further main effect was found for the variable speaker (preceding vowel: $F(3.36) = 55.62$, $MSE = 67361$, $p < .01$; stop closure: $F(3;36) = 6.68$, $MSE = 84492$, $p < .01$); burst ($F(3.36) = 3.29$, $MSE = 69953$, $p < .04$); stop consonant ($F(3;36) =$

2.98, $MSE = 67765$, $p < .05$). The variables speaker and context interacted for the vowel, burst, and stop consonant durations (preceding vowel: $F(3.36) = 17.84$, $MSE = 20666$, $p < .01$; burst ($F(3.36) = 4.22$, $MSE = 65888$, $p < .02$); stop consonant ($F(3;36) = 3.52$, $MSE = 68714$, $p < .03$). Not surprisingly, there is a certain amount of variation between speakers. For example, speaker 4 varied vowel length between clitic, final, and clitic context to a higher degree than speaker 1 (133, 198, 162 ms for speaker 4, and 101, 130, 103 ms for speaker 1, respectively). But all speakers produced longer segments in the final than in the clitic context.

The interaction of speaker, context, and voice was never significant. Tests of simple effects investigated the effect of underlying voicing within each context condition for each speaker separately. The only effect was found for speaker 2, who produced longer vowel preceding voiced as compared to voiceless stops in the clitic context (91 ms and 79 ms, respectively). This single speaker hence produced one of the acoustic cues to the different voicing of the stops in the clitic context. However, in the perception experiment, participants assigned a similar number of voiced ratings to the underlyingly voiced and underlyingly voiceless stops that had been produced by this speaker. This shows that the difference in vowel length alone, while the other segments did not differ in length, was not a sufficient cue for the perception of the voicing contrast.

Two other findings are worth mentioning. The first finding was a higher proportion of items with acoustic energy during closure in the clitic than in the final context. But importantly, voice energy during closure was present in 14% of the underlyingly voiced and 11 % of the underlyingly voiceless stops in the clitic context. The slightly higher proportion for the voiced targets is in the direction as predicted by phonetic theory. But the difference is not huge enough to consider it a reliable cue for voicing. The reason for the higher proportion of energy during closure in the clitic condition as compared to the final context is probably the intervocalic position in the clitic context, opposed to the final stop, which is only preceded, but not followed by a vocalic sound.

A second finding was the following. Comparing the durational patterns obtained for the plural and the clitic context, it turned out that participants produced longer vowels and longer voiceless closures, bursts, and stops in the plural than in the encliticized forms. At first sight, this is surprising, because the plural items, like *raten en helmen* /ra:.tə.nən.hɛl.mən/ ("honeycombs and helmets"), contained more syllables than the clitic items, e.g. *raat en helm* /ra.tən.hɛlm/ ("honeycomb and helmet"). It has been stated in the literature that the length of a syllable decreases with an increasing number of following syllables (Lehiste, 1970; Kohler 1982).

There are several possible explanations for the longer units in plural forms. Firstly, the shorter durations of closure, burst, and stop consonant in the clitic context could be due to the fact that stops are more voiced, because they stand between two vowels. The higher proportion of items with voice energy during closure in the clitic condition had also suggested that view. The voiceless items in the plural context are all voiceless, and therefore have longer closure and burst durations. However, the stops are also in intervocalic position in the plural forms. Moreover, if the items in the clitic context would have been produced with more voicing than the plurals, the vowel preceding the stop consonant should be longer in the clitic condition than in the voiceless plural items. The opposite pattern was found: The vowel segments in the plural context were longer than in the clitic context. On the other hand did the vowel segment not cue voicing to a large extent in the plural context, anyway. The clear contrast between underlyingly voiced and voiceless items was cued by the burst, while the preceding vowel varied only slightly between voiced and voiceless stops (125 ms and 118 ms, respectively). In sum, the difference in the segmental durations is likely to have a different source than the fact that stops in the clitic context were accompanied by more voicing.

A second possibility is that a rhythmic effect causes the pattern: Participants 'rest' on the first syllable in a plural form, which is followed by two weak syllables (e.g., *ponden en olmen*) and they do not do that in the clitic context (e.g., *pond en olm*). This would not only explain the longer closure, burst, and stop duration, but also the longer vowels in the plural as compared to the clitic context.

Thirdly, participants may contrastively pronounce voiced and voiceless stops in the plural context, which is the only context that exhibits both voiced and voiceless stops.

A fourth possibility is to interpret the difference between the two contexts as caused by different acoustic characteristics that resyllabified stops exhibit as compared to stops that do not resyllabify. Dejean de la Bâtie (1993) found a difference in length for stops that occur word-initially and stops that have been resyllabified by encliticization (liaison) in French. She compared word initial /t/s (e.g., *petit tèdre* "little 'tèdre' (=nonword)") with liaison /t/s (e.g., *petit èdre* "little 'èdre' (=nonword)"). Liaison /t/s showed much shorter closure durations and VOT-values than the word-initial /t/s. Both measures taken together indicate that the liaison consonant was uttered faster than the word-initial one. Similarly, a Dutch stop that has been resyllabified in encliticized forms may be produced faster than a stop that has been in the onset of a syllable already lexically in the plural forms.

In the preceding analysis, the number of tokens that were measured in the plural context had been rather small (eight voiced and eight voiceless measurements). Therefore, a second analysis was conducted that included more data points. Of the 11 repetitions of each item that participants had produced in the repeated articulation task in the Second Production Experiment, three were measured. In addition to the items that were tested in the Second Perception Experiment (and analyzed in the previous analysis), the rendering preceding and following that item were measured, resulting in 312 measurements for each preceding vowel, stop closure, release, and stop consonant in the final context, and the same number of measurements in the clitic context, and 48 measurements of each segment in the plural context. It turned out that the mean durations of the segments in the different conditions were only minimally different from the previous analysis. The durational values that were obtained with the three repetitions differed in on average 1.3 ms from the values shown in Table 26. The difference never exceeded 5 ms. The 5 ms difference occurred for the closure of underlyingly voiced stops in final context. In contrast to the former analysis, it was identical to the closure of the underlyingly voiceless stops (140 ms for both voicing conditions). The proportion of items that had voice energy during closure also resembled the values of the previous analysis: A voice bar was present in 13% of the underlyingly voiced stops and 11 % of the underlyingly voiceless stops in the clitic context, in 2% of the underlyingly voiced and 0% of the voiceless stops in the final context, and in 100% of the voiced, but only 4% of the voiceless stops in the plural forms.

Consequently, Anovas over the larger set of measurements revealed the same results as the previous analysis. The only exception was that two interactions that had been found in the previous analysis disappeared: The effect of voicing for the stop closure duration in the final context was not found anymore in the analysis over three repetitions of each item. Furthermore, the effect that had been obtained for speaker 2, who produced longer vowels prreceding underlyingly voiced stops in the clitic context, disappeared. Importantly, as in the previous analysis, did the voice contrast again surface in the plural context, where voiceless stops were accompanied by a burst, while voiced stops lacked it. In the final and the clitic context, the voice contrast was completely neutralized.

In sum, the phonetic measurements revealed that the acoustic aspects of underlyingly voiced and voiceless stops were clearly different in plural forms like *raden* and *raten*. This had been expected, because the stop is in the onset of the second syllable in these forms and therefore no syllable-final devoicing can neutralize the voice contrast. However, this contrast was neutralized in both, the final and the clitic context, where underlyingly voiced and voiceless stops did not differ acoustically. For the final context, this had been clearly expected. The final stops are necessarily in coda position, because no onset follows to which they could attach. Conversely, in a clitic context, like in *raad en*, the final stop of the first word surfaces as the onset of the second syllable. According to the model of phonological encoding, speakers immediately produce the surface syllables. Consequently, the stop in *raad en* does never occupy a syllable-final position, where it could devoice, and should therefore remain voiced. Alternatively to this view, many phonologists (see, e.g., Booij & Rubach,1987; Harris, 1983) suggest that the surface syllable structure is preceded by an abstract syllable structure, which contains syllables for the individual words, for instance, *raad* and *en*. Syllable-final devoicing can neutralize the voice contrast at that level, before the stop resyllabifies to the following onset. The acoustic measurements showed that the contrast between voiced and voiceless stops indeed neutralizes in the clitic context, similar to the contrast for utterance-final stops. These findings parallel the results of the two perception experiments reported above.

SUMMARY AND CONCLUSIONS

Finally, the experimental results are summarized, followed by two sections that discuss the implications of the results for phonological encoding as well as for phonological theory.

Summary of the Results

This thesis studied the role of the syllable in the production of spoken language. When we produce connected speech, word boundaries will not always coincide with syllable boundaries. Weak function words, like pronouns, often fuse with an adjacent word, resulting in cliticized forms. For instance, in the sentence *she gave it*, we usually do not say $(\text{gɛiv})_\sigma(\text{ɪt})_\sigma$, but $(\text{gɛɪ})_\sigma(\text{vɪt})_\sigma$. The same holds for Dutch *kook het* "cook it", which syllabifies $(\text{ko:})_\sigma(\text{kət})_\sigma$. In both examples will the last segment of the verb surface in the syllable onset of the following pronoun. This phenomenon is not at all unusual, because languages universally tend to avoid syllables that start with a vowel.

The only model of phonological encoding in speech production that covers the production of strings of connected speech and cliticized forms is Levelt's (1992, 1993) model. In this model, a segmental spellout procedure provides an ordered set of phonological segments contained in the string of speech that is currently generated, for instance, /k, o:, k, ə, t/ in *kook het.* These segments are then associated with an independently generated prosodic frame. The prosodic frame specifies a string's number of syllables and stress pattern. It has the size of a prosodic word. For *kook het*, there will be one frame, comprising a stressed and an unstressed syllable. Based on Levelt's model, I formulated two predictions with respect to how speakers generate syllables. These two predictions were tested in the two parts of this thesis, respectively.

The first part of the thesis addressed the time course of syllabification. According to the model, syllables are generated at a late point in time during phonological encoding, that is, only after the segments have been spelled out and the prosodic frame has been created. In a series of priming experiments, participants first learned pairs of semantically associated words by heart (e.g., *soep - koken* "soup - to cook") and produced the target (e.g., *koken*) as a response to the visually presented semantic associate (e.g., *soep*). Participants did not only produce the target verb's infinitive forms like *koken*, but also past tense (e.g., *kookte* "cooked") or encliticized forms (e.g., *kook het*). These forms share the initial phonemes, but differ in their syllable structure. The first syllable in the infinitives and encliticized forms of the type *kook het* are short, while past tense forms and encliticized forms of the type *kookt het* ("cook (pl.) it") have a long first syllable.

In addition, participants heard interfering stimuli (IS) like /ko:/ or /le:r/, which were presented at several SOAs with respect to the presentation of the visual cue. The IS were either related to the target (e.g., /ko:/ for *koken*), or unrelated (e.g., /le:/). Previous research had shown that participants react faster when they hear related IS, as compared to unrelated ones, and this effect indeed showed up in all experiments. Furthermore, the IS were either short (e.g., /ko:/, /le:/) or long (e.g., /ko:k, le:r/). In the experiments, I tried to find evidence for two effects that Levelt's model of phonological encoding predicts:

The model predicts that when the interfering syllable is presented at an early SOA - for instance, preceding the presentation of visual cue - only the target's phonological segments, but no syllables have been constructed during phonological encoding. Therefore, the target word's syllable structure should not matter, but only the number of segments that are primed by the syllable. When they hear early IS, participants should react the faster, the more segments are shared by IS and targets, as compared to unrelated IS of corresponding length. I called this the 'segmental overlap effect'.

When the IS are presented at a later SOA - for instance, simultaneously with or after the presentation of the visual cue - the target's syllables should be available. IS that correspond to the target's first syllable should speed up participants' responses more than IS that do not correspond to the syllable, again compared to unrelated IS of corresponding length. This effect was called the 'syllable match effect'.

Alternatively, both effects could co-occur, for instance, when the processing speed differs between participants or items. Under this assumption, they cancel each other for targets with short first syllables. For example, the short stimulus /ko:/ should speed up reactions for the target *koken* (ko:)$_\sigma$(kən)$_\sigma$ because it corresponds to the target's first syllable, while the long stimulus /ko:k/ should speed up reactions because it shares more segments with the target. Reactions to targets with a long first syllable, on the other hand, should be speeded up more efficiently by long IS, according to both the segmental overlap and the syllable match effect.

Participants reacted significantly faster to related than to unrelated IS in all experiments. Moreover, in most experiments they reacted slower when they heard IS (both related and unrelated) that were long than when they heard short ones. Both results show that the experiments tapped into phonological encoding. Furthermore, control experiments excluded that morphological or lexical factors caused the effect. But none of the experiments revealed evidence for an early segmental overlap or a late syllable match effect.

With respect to encliticized targets, short IS speeded up reactions to targets of the type *kook het* more than long IS at a late SOA in Experiment 4. However, this difference, which looked like a late syllable match effect, since the targets' first syllable was short, did not reach significance. Experiments 5a,b tested encliticized forms and varied the targets' syllable structure. While short and long IS speeded up reactions to targets with a short first syllable (e.g., *kook het*) to a similar extent, as predicted for the case that segmental overlap and syllable match effect cooccur, the efficiency of short and long IS was also similar for targets with a long first syllable (e.g., *kookt het*). According to the cooccurring effects, long IS should surpass short IS in prime efficiency for targets with a long first syllable. But since the number of the latter type of targets was rather low, and participants moreover reported that they found these rather difficult to produce, one should not overvalue the results for the latter targets.

Importantly, Experiment 2 showed an interaction that could be interpreted as due to the co-occurrence of segmental overlap and syllable match effect. While the production of infinitives was speeded by related short and related long IS to a similar extent, past tense forms benefitted more from related long than from related short IS. Unexpectedly, this interaction was found at an early SOA and disappeared later, while phonological priming was still obtained at the later SOA. A similar pattern occurred at SOA 150 ms in Experiment 3, but the interaction did not reach significance.

In sum, whether the prediction that syllables are produced late during phonological encoding, is correct, is still an open question. The result obtained in Experiment 2, however, is promising. The different gain participants had from short and long related stimuli did interact with the syllable structure of the target forms. This suggests that syllables are relevant units in speech production. Further research has to investigate the time course of their production more closely.

The second part of this thesis addressed a second prediction of Levelt's model of phonological encoding, which claims that there is only one level of syllabification, namely the surface syllable structure. One could assume that a speaker first syllabifies the individual words to be produced, for instance, *kook* $(\text{ko:k})_\sigma$ and *het* $(\text{ət})_\sigma$. In a next step, these syllables are concatenated, and the consonants that precede a syllable which starts with a vowel will become the onset of that syllable, resulting in an encliticized form like $(\text{ko:})_\sigma(\text{kət})_\sigma$. This account involves two levels of syllable structure: The underlying and abstract lexical syllables $(\text{ko:k})_\sigma$ and $(\text{ət})_\sigma$, and the surface syllables $(\text{ko:})_\sigma$ and $(\text{kət})_\sigma$. In contrast, Levelt's model does not assume an intermediate level of lexical syllables, based on arguments of elegance and economy in processing- why should a speaker produce syllables that never surface? It is much more elegant to assume that segments are not marked for syllable positions, but get their syllable position online during segment-to-frame association. Because only postlexical syllables are created, this syllable position then does not have to be changed anymore. Given the speed at which we produce syllables - about three to five syllables per second - this seems a plausible solution.

To test this claim, I looked for a phenomenon that affects coda consonants. A consonant that surfaces in onset position in encliticized forms (e.g., the /k/ in (kət)$_\sigma$ of *kook het*) should be left unaffected by a regularity that affects only codas. However, if this consonant does show traces of a coda regularity, we would have to conclude that it had been in a coda position before it became an onset. This would constitute evidence for an intermediate level of syllable structure which precedes the surface syllables level. Syllable-final devoicing is a phenomenon of this kind: While Dutch has voiced and voiceless phonemes in syllables onsets (compare the minimal pair *raden* "commissions" and *raten* "honeycombs"), obstruents are voiceless in syllable codas. As a consequence, the singular forms of *raden* and *raten* are the same: (ra:t)$_\sigma$. While the voice contrast clearly survives in the plural forms and is neutralized in the singular forms, encliticized forms are a case for the intermediate level: Under the prediction that surface syllables are the only syllables that are built during phonological encoding, a stop that is voiced in plural forms should also be voiced in encliticized forms, because it is never in a coda position, where it could devoice. For instance, in the encliticized forms *raad en* ("commission and"), where the /d/ surfaces in the onset of the second syllable, it should be voiced: (ra:)$_\sigma$(dən)$_\sigma$. Alternatively to Levelt's model, speakers could first produce the syllables that correspond to the single lexical items, like *raad* and *en*. The syllable-final stop could then devoice before it resyllabifies into the onset of the surface syllable structure.

Two perception experiments tested minimal noun pairs like *raad* and *raat*. The voice contrast was neutralized when the stops occurred at the end of an utterance, and hence also in syllable codas, like in ... *en raat* (' ... and honeycomb'). Participants could not distinguish words like *raad* and *raat* in this context. In plural forms, on the other hand, subjects perceived voiced stops as voiced and voiceless stops as voiceless, because the stops were in onset position. Importantly, the results showed that in encliticized forms, where the stop surfaces in onset position, the voice contrast was nevertheless neutralized. As in the utterance-final cases, participants could not distinguish between *raad* and *raat*. Acoustic measurements on the stimuli of the perception experiments yielded the same results as the perception data. A theory with two levels of syllabification can account for the results by assuming that devoicing in syllable codas applies on the output of the lexical level and thus precedes resyllabification, which links a coda consonant to the onset of a following V-initial syllable. Levelt's model of phonological encoding, on the other hand, cannot readily explain why stops in syllable-onset position obey a constraint for syllable codas in the encliticized forms, because according to the model the stop is never in coda position during the process of syllabification.

Implications for Phonological Theory

The results of the second part of this thesis, which showed syllable-final devoicing of stops that surface in onset position of encliticized forms, provide empirical data that are relevant for some issues in phonological theory. One ongoing debate is whether final extraprosodicity (e.g., Itô, 1986; Rice, 1990) or resyllabification should be preferred in theories of syllabification (for a discussion see Hall, 1994). Many phonological theories (e.g., Harris, 1983; Rubach & Booij, 1990) assume resyllabification not only between, but also within words, because there are cycles of syllabification throughout the morphological derivation of a word. When, for instance, a vowel-initial suffix (e.g., the Dutch plural suffix *-en*) is added to a stem, this stem will already be syllabified. If the stem has a final coda consonant, like in the morpheme *raad*, this consonant will delink from the coda of the first syllable and simultaneously reassociate to the onset position of the following syllable, which is provided by the plural suffix.

According to final extraprosodicity, on the other hand, the final consonant of a root morpheme is considered to be extraprosodic, that is, the consonant is invisible for syllabification. The morpheme *raad* will syllabify $(\mathrm{ra{:}})_\sigma$<d>. When a vowel-initial suffix is added, like the plural suffix *-en*, the final consonant of *raad*, which does not yet belong to any syllable, can associate with the following onset, as in $(\mathrm{ra{:}})_\sigma(\mathrm{d\partial n})_\sigma$. In contrast to resyllabification, the consonant is assigned to a syllable position only once.

Theories that assume final extraprosodicity differ with respect to the point when extraprosodicity turns off and the consonant participates in syllabification. According to Itô (1986), extraprosodicity universally turns off at word level, i.e., final consonants become visible and must be syllabified after the morphological derivation of a word is completed. However, there are cases where the word-final consonant surfaces in onset position of a following word, like in the encliticized forms. To account for these, Itô has to assume postlexical resyllabification. Itô hence needs two different mechanisms, extraprosodicity and postlexical resyllabification. Since she allows the latter, she can explain the final devoicing in Dutch encliticized forms. In contrast to that, Rice (1990) claims that final extraprosodicity remains active beyond word level. As a consequence, she can treat within- and between-word syllabification by the same mechanism and avoids any resyllabification. She prefers final extraprosodicity, because it is only structure-building, while resyllabification is structure-changing: Resyllabification involves the delinking of the consonant from the coda position, which in fact destroys the first syllable structure before a new one is built. Final extraprosodicity seems to be more restricted, since it only creates new structure, and is therefore to be preferred. However, to account for final devoicing in Dutch encliticized forms, the obstruent has to occupy a coda position at some point, because underlyingly voiced stops surface voiceless in the postlexical onset position. Dutch final devoicing in encliticized forms adds hence another case in favor of resyllabification to the debate of resyllabification and final extraprosodicity.

A further option would be to avoid resyllabification by assuming that the consonant in encliticized forms like *raad en* is in fact ambisyllabic, belonging to both onset of the second and coda of the first syllable. Since it is still associated with a coda position, it could devoice (see also the discussion in Booij, to appear). However, ambisyllabicity in Dutch in fact blocks devoicing. The plural of *pad* [pɑt] "path", for instance, is *padden* [pɑdən] with an ambisyllabic voiced stop. Syllable-final devoicing applies in the singular, but not in the plural form. The notion of ambisyllabicity thus cannot replace resyllabification in Dutch.

The data in the second part of the thesis are of interest for the question whether the phonological component includes intermediate levels of syllabification. In the traditional framework of Lexical Phonology, (e.g., Kiparsky, 1985; Kaisse & Shaw, 1985; Booij & Rubach, 1987) one can account for the present findings (see also Booij, 1996). Lexical Phonology distinguishes a lexical and a postlexical component, where the output of the former provides the input to the latter component. Final devoicing applies at the end of the lexical level on the single lexical units, preceding the postlexical rule component. In contrast to that, the data of Part 2 are problematic for monostratal phonological theories that do not distinguish lexical and postlexical components. And the data furthermore show that phonological theories that replace the traditional phonological rules and derivations by other means have to make additional assumptions to account for the data (see Booij, to appear, for a detailed discussion). An example for a phonological framework that is not rule-based is Optimality Theory (McCarthy & Prince, 1993; Prince & Smolensky, i. pr.). In this theory, the underlying representation of, for instance, a word is paired with a whole set of candidates for the surface structure of this word. Underlying representations are unsyllabified segmental strings. The output candidates are fully prosodified, including stress information, epenthetic segments, syllable structure, and so forth. Which of those candidates will actually surface is evaluated by a set of wellformedness constraints. Some of these constraints refer to syllable constituents, for example, the constraint that states that syllables have to have an onset. Constraints may be violated. Normally, each candidate will violate some of the constraints. Importantly, the constraints are ranked on a hierarchy of priority. Candidate forms that violate the highest-ranked constraint are not considered anymore as possible surface forms. The candidates that are in harmony with the highest-ranked constraint are evaluated based on the constraint that is ranked one level down in the hierarchy. This procedure is repeated until only one candidate is left. This candidate form, the 'optimal candidate' will be the surface form of an underlying representation. As mentioned, this candidate may still violate some constraints, but these have to be lower ranked than the constraint that caused the elimination of the last alternative candidate to the optimal candidate form. Lower ranked constraints do not matter anymore.

How could Optimality Theory account for final devoicing in encliticized forms like *raad en*? One could imagine that the following output candidates (among many others) are included in the candidate set: $(\text{ra:})_\sigma(\text{dən})_\sigma$, $(\text{ra:})_\sigma(\text{tən})_\sigma$, $(\text{ra:d})_\sigma(\text{ən})_\sigma$, $(\text{ra:t})_\sigma(\text{ən})_\sigma$. The results of Part 2 showed that the second candidate is the one that surfaces in Dutch, hence this candidate should be the optimal one. To rule out the latter two candidates with schwa-initial syllables, one could think of a highly-ranked constraint that prohibits schwa-initial syllables. Only the first two candidates obey this constraint. For final devoicing we need a constraint that rules out candidates with a voiced obstruent in coda position. This constraint, if rather highly ranked, will rule out all candidate forms with syllables ending in voiced final stops or fricatives. However, this constraint will rule out neither of the two remaining candidates, since they both do not have a voiced syllable-final obstruent. The crucial point is that the constraint ranking that rules out voiced codas has to apply at word level, before the syllabification across word boundaries is considered.

Postlexical phonological phenomena have not received much attention within the young framework of Optimality Theory. At the moment, the only option to account for the facts seems to be to allow for two levels of constraint evaluation. First, surface candidates are evaluated at a lexical level, and the constraint that regulates syllable-final devoicing has to rank high within that level. The output of the lexical level is then further evaluated by a set of constraints at a postlexical level that rule out candidates that are not encliticized. Importantly, these two levels of constraint evaluation have to be serially ordered, because postlexical evaluation operates on the basis of the output of the lexical level.

A further ongoing discussion concerns the relation between phonetics and phonology. Are contrasts that are neutralized phonologically - like the neutralization of voicing in syllable-final position - also completely neutralized phonetically? I found complete phonetic neutralization of the contrast in Dutch minimal pairs. Jongman et al. (1992) had also found complete neutralization of preceding vowel duration and consonant closure duration for final underlyingly voiced and voiceless consonants in Dutch minimal pairs. As discussed in Part 2, the experimental results reported in the literature on other languages are not clear. Some studies found complete, others incomplete phonetic neutralization. The relation of phonology and phonetic implementation definitely needs further research. To investigate how the voicing contrast and its neutralization is phonetically implemented, Dutch offers a good alternative to German, which has mostly been studied for that purpose, because it has a greater number of minimal pairs that consist of known words of the same word class, probably due to a higher number of monosyllabic nouns in Dutch.

Implications for Phonological Encoding

With respect to research on phonological encoding in general, the experiments in the first part showed that consistent phonological priming effects occur with auditory IS that are spoken syllables. The IS in previous research had always been words. Furthermore, the semantic-associate learning task, which had been used in an implicit priming task before (Meyer, 1988, 1990, 1991), proved its worth in an online priming paradigm. This is useful, because the methods to elicit utterances in production research are limited. Semantic-associate learning has the further advantage that it can be used to elicit abstract targets and targets that consist of several words.

There are several possibilities to adapt Levelt's model of phonological encoding such that it can explain the findings of Part 2, which showed that the stop surfacing in onset position of encliticized forms like *raad en* is devoiced. First, one could try to leave the model as it is, but explain the findings differently and argue that the syllable is not the domain of devoicing at all. The syllable has not always been the most popular candidate to explain final devoicing. Since it was not available as a unit for most phonologists in the early generative tradition, they treated final devoicing as a morpheme-final phenomenon (e.g., Wurzel, 1970 for German). Following this account, final devoicing would apply at the end of morphemes, if these are words or members of compounds, or if a non-vocalic element follows the obstruent. This rule, however, does not refer to a natural phonological class, as phonological rules normally do. There are no other phonological processes that need to divide suffixes into a consonant-initial and a vowel-initial group. The rule that takes the syllable coda as a domain is much simpler and refers to an environment that is needed anyhow, for instance, to explain the distribution of the velar nasal that is only licensed in syllable codas, or of the glottal fricative /h/ that is not allowed in Dutch codas. Moreover, the morpheme-final account cannot explain why final devoicing is so productive even when speakers produce, for instance, nonwords, or acronyms, or foreign names like *Sy*[t.n]*ey* or *Cambo*[t.j]*a* (Booij, 1977). The assumption of morpheme-final devoicing is thus not desirable. Devoicing applies syllable-finally. Alternatives to resyllabification like final extraprosodicity or ambisyllabicity have already been discussed and rejected in the preceding section.

A different possibility is to avoid devoicing completely in the production model. One could argue that all morphologically derived or inflected words - at least the suffixed ones - are stored. Take, for instance, the word *voogd* ("guardian"). When the suffix starts with a vowel, the forms will be stored with a voiced stop, e.g., *voogd-en* $(\text{vo:x})_\sigma(\text{dən})_\sigma$ "guardians", *voogd-es* $(\text{vo:x})_\sigma$ $(\text{dəs})_\sigma$ "guardian (Genitive)", *voogd-ij* $(\text{vo:x})_\sigma(\text{dɛɪ})_\sigma$ "guardianship". For forms where the suffix starts with a consonant (and for the exceptional suffix *-achtig*), the stop is stored as voiceless, e.g., *voogd-loos* $(\text{vo:xt})_\sigma(\text{lo:s})_\sigma$ "guardianless". Also word-final obstruents would be voiceless already in the lexicon. Under these assumptions, the model correctly predicts voiceless stops in encliticized forms.

Unfortunately, we do not know much about how morphologically complex words are produced. While irregular inflected forms like the suppletive forms in the verbal paradigms of *go* (*went - gone*) or *be* (*am - is - are*) are likely to be stored as a unit in the lexicon, things look differently for regular forms. Stemberger and MacWhinney (1986) argued for the storage of high-frequency regular inflected forms in the lexicon, because they obtained effects based on the specific frequency of an inflected form. High-frequent inflected items showed fewer phonological errors than low frequent inflected items in a corpus of speech errors. This held for both regular and irregular inflections. Furthermore, subjects who had to produce past tense verb forms in an experiment made significantly more no-marking errors - i.e., omissions of the past tense morpheme - for low-frequent regular inflected forms. Therefore the authors suggest that high-frequent regular inflected forms are stored, while lower-frequent ones are generated by rule.

But some arguments speak against the full storage of morphologically complex forms. A phenomenon often mentioned in speech error literature is affix stranding. While words or stems exchange in speech errors, inflectional or derivational affixes stay at their position in the utterance, see the examples in Table 27. Moreover, affixes accommodate to the new phonological contexts that result from speech errors (Garrett, 1980, Stemberger, 1985), which also indicates an analysis into morphemes. Since Stemberger and MacWhinney found affix shifts and morphological accommodations not only for low-frequent, but to a similar extent also for high-frequent regular inflected forms, they conclude that the high-frequent forms are both, stored as a whole - what explains the frequency effects for the no-marking errors, and analyzed into their morphemes - what explains the error pattern in the corpus.

error type	error	target
Inflectional Affix Stranding:	`words of rule formation`	`rules of word formation`
New Derivational Affix:	`the introducting of the ...`	`the introduction of the ...`
Affix Shift:	`if she wantø to comes here`	`if she wants to comeø here`
Morphological Accomodation:	`bunnie/z/ don't eat steak`	`steak/s/ don't eat bunny`

Table 27. Morphological Speech Errors

Speech errors provide more arguments in favor of the production of morphologically complex words by rule. Inflectional or derivational affixes may only be moved or shifted, but do not frequently exchange, while words and stems do not move or shift very often (Garrett, 1988). Similarly, regular inflections may shift to an adjacent word as in *tell-us-ing* for *telling us* (Stemberger & MacWhinney, 1986), while nonmorphemic word-final material seldomly shifts. The fact that some errors result in nonexisting, but possible derived forms points against the full storage hypothesis, too. Table 27 shows some relevant errors, taken from Fromkin (1993).

A further argument against the full storage of derived forms as unanalyzed units is that morphological information is crucial for syllabification, as discussed in the introduction to this thesis, because some non-cohering suffixes, for instance *-achtig*, and all prefixes are separate domains for syllabification. However, one could posit a model (like Jackendoff, 1975) that stores morphologically complex forms as a sequence of morphemes, such that this information is maintained. To account for the productivity of final devoicing, one has moreover to assume that speakers know the contexts in which final devoicing occurs: Devoiced obstruents could be split into two groups that result from different devoicing mechanisms. Syllable-final obstruents are voiceless, and in addition all word-final obstruents are voiceless, too. The latter addition covers also the voiceless stops in the encliticized forms.

Alternatively to storing morphologically complex words, they could be constructed. The two groups of obstruents could then devoice in the course of the encoding process. An obstruent that precedes a lexical word boundary would have to be marked early in phonological encoding. Two scenarios are possible. In the first one, the word-final obstruent is immediately devoiced, while 'ordinary' syllable-final devoicing waits until the surface syllables are created. This implies two levels of devoicing during production. Alternatively, the marker for a following word boundary is carried through the segmental spellout procedure and the segment-to-frame association process, until one process devoices both, syllable-final obstruents and those with the word boundary marker. Crucially, this mark may only be set after the cohering V-initial suffixes have been attached - otherwise obstruents would also devoice, for instance, in plural forms. Since non-cohering suffixes are prosodic words on their own and therefore separate domains of syllabification, the obstruents preceding them automatically stay in coda position, such that they devoice anyway - due to a word-final boundary marker or a coda devoicing mechanism. In this solution, higher level information, i.e., the position of lexical boundaries, has to be inherited through the phonological production process.

A solution that changes the model in a more substantial way, since it attacks one of its major claims, is to allow for resyllabification, that is, to assume two levels of syllable structure during encoding. In fact, Dutch final devoicing in encliticized forms is not the only phenomenon that refers to an underlying lexical syllable structure. Other languages have rules that affect a coda element before this element resyllabifies into the following onset as well. In many Spanish dialects (Harris, 1983), for instance, nasals are velarized and the fricative /s/ is realized as /h/ in coda positions, see (10a,b):

(10)	a.	Ra.mó[ŋ]	Ramon	"Ramon (name)"
	b.	tie.ne[h]	tienes	"you have"

In resyllabified strings of connected speech, where the nasal or the fricative occupies an onset position, they nevertheless undergo the coda rules, see (11a,b). As the devoiced final stops in Dutch encliticized forms, the elements must have been in coda position before resyllabification.

(11)	a.	Ra.mó.[ŋ] en.tró	Ramón entró	"Ramon enters"
	b.	tie.ne.[h]es.pa.cio	tienes espacio	"you have room"

Korean has a rule that makes the fricatives /s/ and the glottalized /s'/ become stops syllable-finally (see Hall, 1994). This rule again precedes postlexical resyllabification, as example (12) shows. In example (12c), in addition to resyllabification a rule has applied that voices intervocalic nonaspirated, nonglottalic stops.[17]

(12)	a.	[ot]	/os/	"clothes"
	b.	[o.ʃi]	/os-i/	"clothes (subj.)"
	c.	[o.di.pe]	/os ip-e/	"cloth put.on"

The English /l/-allophones also serve as an example: Spoken in isolation, the word *fill* will end in a dark /l/, which is the variant that occurs syllable-finally. The sentence *Fill every cup*, spoken at normal tempo, however, will contain the clear variant that occurs in syllable onsets, due to resyllabification (Gussenhoven, 1986).

In Québécois French, vowels are lax in closed syllables and diphthongized in open syllable. For instance, the masculine form of the adjective *petit* is (pt^{s}ij)$_{\sigma}$, the feminine form is *petite* (pt^{s}ɪt)$_{\sigma}$. When the feminine form precedes a V-initial word, e.g. *petite amie*, the final /t/ resyllabifies: (pt^{s}ɪ)$_{\sigma}$(tɑ)$_{\sigma}$(mij)$_{\sigma}$ Although the first syllable is open now, the vowel is not diphthongized, but remains lax.

The different rules mentioned above all share the same feature: Like final devoicing in Dutch encliticized forms, they affect elements in coda position before these elements resyllabify into the onset of the following syllable whose onset position is empty. A model that avoids resyllabification has to assume that Spanish has two rules of nasal velarization, a syllable-final and a later word-final rule, and two rules of fricative aspiration, again a syllable-final and the later word-final rule, and the same would have to hold for the other phenomena in the different languages. This appears ad hoc compared to an account that unifies the syllable- and word-final rules of each language to one syllable-final rule and incorporates resyllabification, which reflects the general tendency of languages to avoid syllables without onsets.

Resyllabification was included in earlier versions of Levelt's model, which assumed that syllabification involves two steps. First, single lexical items were syllabified. In a next step, these were combined into one prosodic word, which involved resyllabification (Levelt, 1989: 406 ff). In the recent version of Levelt's model (1992, 1993), resyllabification is abandoned based on arguments of elegance and economy in processing. On the other hand, one can argue that resyllabification does not take that much processing effort. As a matter of fact, it is

[17] I would like to thank Soonja Choi for checking the Korean examples.

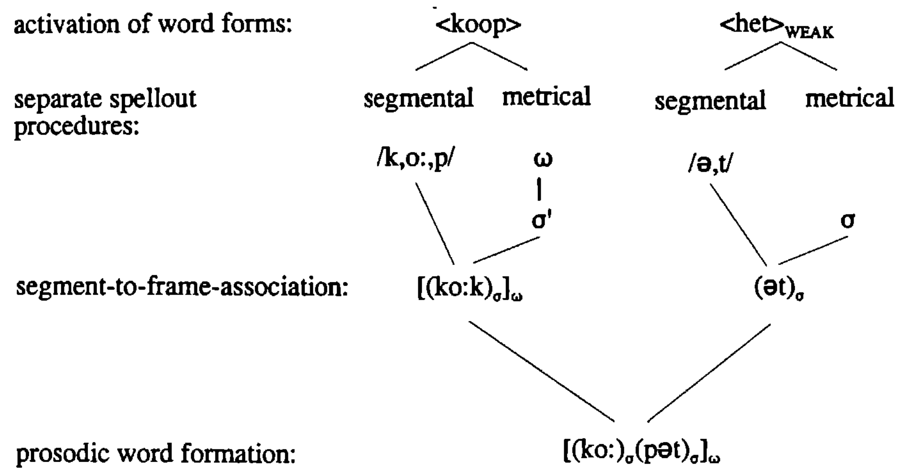


Figure 22. Resyllabification in the Phonological Encoding Component

only one segment at the edge of a lexical unit whose syllable affiliation alternates and that has to be linked from the coda of one to the onset of the following syllable. Figure 22 shows how resyllabification could work in phonological encoding. Lexical words, e.g., *kook* ("cook") and *het* ("it") are each syllabified separately. Then the syllables of the two words combine to one prosodic word and the final stop of the first word resyllabifies to the onset of the following word. The resulting postlexical syllables are the input for the articulatory component.

In sum, while other models of phonological encoding cannot explain the production of encliticized forms at all, Levelt's model can handle the data by some additional assumptions which may or may not involve resyllabification. To avoid resyllabification, either lexical boundaries have to be retained until a late point in phonological encoding, or the word-final voice neutralization is stored with morphologically complex items. Both options require in addition a syllable-final mechanism for devoicing in foreign words, names, and abbreviations. Alternatively, if one allows for postlexical resyllabification, all devoiced obstruents have a common origin in the prohibition of voiced obstruents in syllable codas. Which account is correct cannot be decided on the basis of the present experiments. But because of the phonological arguments presented above, I prefer to allow resyllabification during the production of syllable structure in connected speech.

REFERENCES

Anderson, S., & Port, R. (1994). Evidence for syllable structure, stress and juncture from segmental durations. *Journal of Phonetics*, *22*, 283-315.

Anderson, S. R. (1985). *Phonology in the twentieth century. Theories of rules and theories of representations.* Chicago & London: University of Chicago Press.

Baars, B. J., Motley, M. T., & MacKay, D. (1975). Output editing for lexical status from artificially eicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour*, *14*, 382-391.

Balota, D. A. , & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89-106.

Baumann, M. (in press). The production of syllables in cliticized forms. In A. Dainora, R. Hemphill, B. Luka, B. Need, & S. Pargman (Eds.), *CLS 31-II: Papers from the Parasession on Clitics*, Chicago: Chicago Linguistic Society.

Berendsen, E. (1986). *The phonology of cliticization.* Dordrecht: Foris.

Blaauw, E. (1995). *On the perceptual classification of spontaneous and read speech.* (Doctoral Dissertation, Utrecht University. Published by LEd & OTS, Utrecht).

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 206-244). Oxford: Basil Blackwell.

Bock, K., & Levelt, W. J. M. (1994). Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Booij, G. (1977). *Dutch morphology. A study of word formation in Generative Grammar.* Lisse / Dordrecht: de Ridder / Foris.

Booij, G. (1985). Coordination reduction in complex words: a case for prosodic phonology. In H. van der Hulst, & N. Smith (Eds.), *Advances in non-linear phonology* (pp. 143-160). Dordrecht: Foris.

Booij, G. (1995). *The phonology of Dutch.* Oxford: Clarendon Press.

Booij, G. (1996). Cliticization as prosodic integration: the case of Dutch. *The Linguistic Review*,.

Booij, G. (to appear). Lexical Phonology and the derivational residue. To appear in: J. Durand & B. Laks (Eds.), *Proceedings of the Conference on Models & Trends in Phonology, Royaumont, June 1995.*

Booij, G., & Rubach, J. (1987). Postcyclic versus postlexical rules in Lexical Phonology. *Linguistic Inquiry*, *19*, 1-44.

Boyce, S. E., Krakow, R. A., & Bell-Berti, F. (1991). Phonological underspecification and speech motor organisation. *Phonology*, *8*, 219-236.

Boyce, S. E., Krakow, R. A., Bell-Berti, F., & Gelfer, C. E. (1990). Converging sources of evidence for dissecting articulatory movements into core gestures. *Journal of Phonetics*, *18*, 173-188.

Bradley, D. C., Sánchez-Casas, R. M., & García-Albea, J. E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes*, *8*, 197-233.

Brockhaus, W. G. (1989). Final Devoicing and Neutralization. *UCL Working Papers in Linguistics 3*, 303-330.

Browman, C. P. (1978). Tip of the tongue and slip of the ear. Implications for language processing. *UCLA Working Papers in Phonetics*, *42*, 1-149.

Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An overview. *Phonetica*, *49*, 155-180.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*, 204-223.

Brown, R., & McNeill, D. (1966). The "tip-of-the-tongue" phenomenon. *Journal of Verbal Learning and Verbal Behaviour*, *5*, 325-337.

Bruck, M., Treiman, R., & Caracolas, M. (1995). Role of the syllable in the processing of spoken English: Evidence from a nonword comparison task. *Journal of Experimental Psychology: Human Perception and Performance*, *21/3*, 469-479.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language Production Vol. 2: Development, writing and other language phenomena* (pp. 257-294). London: Academic Press.

Cain, H. (1986). *A lexicon of foreign loan-words in the Samoan language*. Köln, Wien: Böhlau.

Cañas, J. J., & Bajo, M. T. (1994). Strategic associative priming in the lexical decision task. *The Quarterly Journal of Experimental Psychology*, *47A/2*, 383-405.

Catford, J. C. (1992). *A practical introduction to phonetics*. (4th ed.). Oxford: Clarendon.

Charles-Luce, J. (1985). Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics*, *13*, 309-324.

Chomsky, N., & Halle, M. (1969). *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, *2*, 225-252.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology 1: Between the grammar and the physics of speech* (pp. 283-333). Cambridge et al.: Cambridge University Press.

Clements, G. N., & Keyser, S. J. (1983). *CV phonology: A generative theory of the syllable*. Cambridge et al.: Cambridge University Press.

Cooper, W. G., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Crompton, A. (1982). Syllables and segments in speech production. In A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 663-716). Berlin: Mouton.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustic Society of America*, *83/4*, 1553-1573.

Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustic Society of America*, *88/1*, 101-112.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385-400.

Dejean de la Bâtie, B. (1993). *Word boundary ambiguity in spoken French*. (Unpublished Doctoral Dissertation). Victoria: Monash University.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, *93/3*, 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests and predictions of a connectionist model. *Journal of Memory and Language*, *27*, 124-142.

Dell, Gary S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287-314.

Dinssen, D. A. (1984). A re-examination of phonological neutralization. *Journal of Linguistics, 21*, 265-279.

Dinssen, D. A., & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics, 12*, 46-60.

Drews, E., & Zwitserlood, P. (in press). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance.*

Dyer, F. N. (1971). A comparison of chromatic and achromatic versions of the Stroop Color-Word Test. *Psychonomic Science, 22*, 235-237.

Ferreira, F. (1991). Effects of length and sylntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30*, 210-233.

Fourakis, M., & Iverson, G. K. (1984). On the 'incomplete netralization' of German final obstruents. *Phonetica, 41*, 140-149.

Fowler, C. A., Treiman, R., & Gross, J. (1993). The structure of English syllables and polysyllables. *Journal of Memory and Language, 32*, 115-140.

Fromkin, V. A. (1973). Introduction and appendix. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 11-45, 243-269). The Hague: Mouton.

Fromkin, V. A. (1993). Speech production. In J. Berko Gleason, & N. Bernstein Ratner (Eds.), *Psycholinguistics* (pp. 272-300). Fort Worth et al.: Hartcourt Brace College Publishers.

Fujimura, O., & Lovins, J. B. (1978). Syllables as concatenative phonetic units. In A. Bell, & J. B. Hooper (Eds.), *Syllables and segments* (pp. 107-130). Amsterdam: North-Holland.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation: Vol. 9* (pp. 133-177). New York: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1. Speech and Talk* (pp. 177-220). London: Academic Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey. Vol. III. Language: Psychological and biological aspects* (pp. 69-96). Cambridge et al.: Cambridge University Press.

Gerken, L. A. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language, 30*, 431-451.

Gerken, L. A. (1994). Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions. *Journal of Memory and Language, 33*, 19-38.

Gerken, L. A., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure; 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition, 51*, 237-265.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8/6*, 875-894.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of Picture-Word Interference. *Journal of Experimental Psychology: Human Perception and Performance, 10/5*, 640-654.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming on spoken word recognition: The role of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18/6*, 1211-1238.

Goldsmith, J. A. (1990). *Autosegmental and metrical theory.* Oxford: Basil Blackwell.

Groot, A. M. B. de (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology, 36A*, 253-280.

Groot, A. M. B. de (1985). Word-context effects in word naming and lexical decision. *Quarterly Journal of Experimental Psychology, 37A*, 281-297.

Gussenhoven, C. (1980). Review of: Mieke Trommelen & Wim Zonneveld: Inleiding in the generatieve fonologie (1979) Muiderberg: Dick Coutinho. *Gramma, 4*, 175-183.

Gussenhoven, C. (1985). Over de fonologie van Nederlandse clitica. *Spektator - Tijdschrift voor Neerlandistiek, 15/3*, 180-200.

Gussenhoven, C. (1986). English plosive allophones and ambisyllabicity. *Gramma, 10*, 119-141.

Hall, T. A. (1992). *Syllable structure and syllable-related processes in German.* Tübingen: Niemeyer.

Hall, T. A. (1994). Extrasyllabicity and resyllabification. In R. Wiese (Ed.), *Recent Developments in Lexical Phonology. (=Theorie des Lexikons, Arbeiten des SFB 282, Vol. 56)* (pp. 151-166). Düsseldorf & Wuppertal: University of Düsseldorf & University of Wuppertal.

Harris, J. W. (1983). *Syllable structure and stress in Spanish. A nonlinear analysis.* Cambridge, Mass.: MIT Press.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry, 20*, 253-306.

Hays, W. L. (1989). *Statistics.* (4th ed.). New York et al.: Holt, Reinhart & Winston.

Heyer, K. den, Briand, K., & Dannenberg, G. L. (1983). Strategic factors in a lexical-deicision task: Evidence for automatic and attention-driven processes. *Memory & Cognition, 11/4*, 374-381.

Hirsh-Pasek, K. et al. (1987). Clauses are perceptual units for prelinguistic infants. *Cognition, 26*, 269-286.

Hooper, J. B. (1972). The syllable in phonological theory. *Language, 47/1*, 133-140.

Hulst, H. van der (1984). *Syllable structure and stress in Dutch.* Dordrecht: Foris.

Hyman, L. (1985). *A theory of phonological weight.* Dordrecht: Foris.

Itô, J. (1986). *Syllable theory in prosodic phonology.* (Doctoral Dissertation) Amherst, MA: University of Massachussets.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language, 51*, 639-671.

Jassem, W., & Richter, L. (1989). Neutralization of voicing in polish obstruents. *Journal of Phonetics, 17*, 317-325.

Jescheniak, J.-D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20/4*, 824-843.

Jongenburger, W., & Heuven, V. van (1991). The distribution of (word initial) glottal stop in Dutch. In F. Drijkoningen, & A. van Kemenade (Eds.), *Linguistics in the Netherlands 1991* (pp. 101-110). Amsterdam: Benjamins.

Jongman, A., Sereno, J. A., Raaimakers, M., & Lahiri, A. (1992). The phonological representation of [voice] in speech perception. *Language & Speech, 35/1-2*, 137-152.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development. Models, research, implications* (pp. 17-64). Timonium, MD: York Press.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Preference for the predominant stress patterns of English words. *Child Development, 64*, 675-687.

Juzcsyk, P. W. et al. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24*, 252-293.

Kahlen-Halstenbach, B. (1990). Zur psychologischen Realität der Auslautverhärtung im Deutschen. *Zeitschrift für phonetische Sprachwissenschaft und Kommunikationsforschung, 43/5*, 645-655.

Kahn, D. (1976). *Syllable-based generalizations in English phonology.* (Cambridge, Mass. Doctoral Dissertation. Reproduced by Indiana University Linguistics Club).

Kaisse, E., & Shaw, P. A. (1985). On the theory of Lexical Phonology. *Phonology Yearbook, 2*, 1-30.

Keefe, D. E., & Neely, J. H. (1990). Semantic priming in the pronunciation task: The role of prospective prime-generated expectancies. *Memory & Cognition, 18/3*, 289-298.

Kempen, G. & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*, 201-258.

Kenstowicz, M. (1994). *Phonology in generative grammar.* Cambridge & Oxford: Basil Blackwell.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech.* San Diego: Singular Publishing Group.

Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook, 2*, 83-138.

Kohler, K. J. (1982). *Rhythmus im Deutschen.* (Arbeitsberichte). Kiel: Institut für Phonetik.

Kohler, K. J. (1986). Invariance and variability in speech timing: From utterance to segment in German. In J. S. Perkell, & D. H. Klatt (Eds.), *Invariability and variabiliy in speech processes* (pp. 268-299). Hillsdale, NJ: Lawrence Erlbaum.

Kohler, K. J. (1994). Glottal stops and glottalization in German. *Phonetica, 51*, 38-51.

Kooij, J. G. (1980). Morpheme boundaries and syllable boundaries: A case for natural phonology. In W. Zonneveld, F. van Coetsem, & O. W. Robinson (Eds.), *Dutch Studies. Vol. 4. Studies in Dutch phonology* (pp. 61-78). The Hague: Martinus Nijhoff.

Koopmans-van Beinum, F. J. (1982). Akoestische en perceptieve aspecten van klinkercontrast-reductie en de rol van de fonologie. *Spektator - Tijdschrift voor Neerlandistiek, 11*, 284-294.

Koopmans-van Beinum, F. J. (1994). What's in a Schwa? *Phonetica, 51*, 68-79.

Krakow, R. A. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures.* (Unpublished Doctoral Dissertation). Ann Arbor: Yale University.

Kuijpers, C. T. L. (1993). Temporal aspects of the voiced-voiceless distinction in speech development of young Dutch children. *Journal of Phonetics, 21*, 313-327.

Kuijpers, C. T. L. (1993). *Temporal coordination in speech development.* (Unpublished Doctoral Dissertation). Amsterdam: University of Amsterdam.

La Heij, W., et al. (1990). Orthographic facilitation and categorial interference in a word-translation variant of the Stroop task. *Canadian Journal of Psychology, 44*, 76-83.

La Heij, W., Starreveld, P. A., & Steehouwer, L. C. (1993). Semantic interference and orthographic facilitation in definition naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19/2*, 352-368.

Ladefoged, P. (1993). *A course in phonetics.* (3rd ed.). Fort Worth: Harcourt Brace Jovanovich College Publishers.

Lahiri, A., Jongman, A., & Sereno, J. A. (1990). The pronominal clitic [dər] in Dutch: A theoretical and experimental approach. *Yearbook of Morphology, 3*, 115-127.

Lass, R. (1984). *Phonology. An introduction to basic concepts.* Cambridge et al.: Cambridge University Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America, 54*, 1228-1234.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation.* Cambridge et al.: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42*, 1-22.

Levelt, W. J. M. (1993). Timing in speech production: With special reference to word form encoding. *Annals of the New York Academy of Sciences, 682*, 283-295.

Levelt, W. J. M. et al. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98/1*, 122-142.

Levelt, W. J. M., & Wheeldon, L. R. (1994). Do speakers have access to a mental syllabary? *Cognition, 50*, 239-269.

Liberman, I. Y., Shankweiler, D., Fischer, F. W. , & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*, 201-212.

Lombardi, L. (1995). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory, 13*, 39-74.

Macken, M. A. (1992). Where's phonology? In Ferguson, C. A., Menn, L., Stoel-Gammon, C. (Ed.), *Phonological development. Models, research, implications* (pp. 249-269). Timonium, MD: York Press.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109/2*, 163-203.

Maddieson, I. (1984). Phonetic cues to syllabification. *UCLA Working Papers in Phonetics, 59*, 85-101.

McCarthy, J. J., & Prince, A. S. (1993). *Prosodic Morphology I: Constraint interaction and satisfaction.* (Ms.). University of Massachusetts, Amherst and Rutgers University.

McQueen, J. M., & Cox, E. (1995). The use of phonotactic constraints in the segmentation of Dutch. *Proceedings of Eurospeech, 1995, Madrid, 18-21 September*, 1707-1710.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behaviour, 20*, 298-305.

Meijer, P. J. A. (1994). *Phonological encoding: The role of suprasegmental structures.* (Unpublished Doctoral Dissertation). Nijmegen: University of Nijmegen.

Meyer, A. S. (1988). *Phonological encoding in language production. A priming study.* (Unpublished Doctoral Dissertation). Nijmegen: University of Nijmegen.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29*, 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: The phonological encoding inside a syllable. *Journal of Memory and Language, 30*, 69-89.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition, 42*, 181-211.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 1146-1160.

Morais, J. (1985). Literacy and the awareness of the units of speech: implications for research on the units of perception. *Linguistics, 23*, 707-721.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323-331.

Mowrey, R. A., & MacKay, I. R. A. (1990). Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America, 88*, 1299-1312.

Neely, J. H., Keefe, D. E., & Ross, K. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 1003-1019.

Neijt, A. (1984). Clitics in arboreal phonlogy. In H. van der Hulst, & N. Smith (Eds.), *Advances in nonlinear phonology* (pp. 179-192). Dordrecht: Foris.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

Nittrouer, S., Munhall, K., Kelso, J. A. S., Tuller, B., & Harris, K. S. (1988). Patterns of interarticulator phasing and their relation to linguistic structure. *Journal of the Acoustic Society of America, 84/5*, 1653-1661.

Nooteboom, S. G. (1969). The tongue slips into patterns. In A. G. Sciarone, A. J. van Essen, & A. A. van Raad (Eds.), *Nomen: Leyden Studies in Linguistics and Phonetics* (pp. 114-132). The Hague: Mouton.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America, 54*, 1235-1247.

Oller, D. K., & Lynch, M. P. (1992). Infants vocalizations and innovations in infraphonology: Toward a broader theory of development and disorders. In Ferguson, C. A., Menn, L., Stoel-Gammon, C. (Ed.), *Phonological development. Models, research, implications* (pp. 509-536). Timonium, MD: York Press.

Pallier, C., Sebastian-Gallés, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language, 32*, 373-389.

Pierrehumbert, J., & Nair, R. (1995). Word games and syllable structure. *Language & Speech, 38/1*, 77-114.

Piroth, H. G., Schiefer, L., Janker, P. M., & Johne, B. (1991). Evidence for final devoicing in German? An experimental investigation. *Proceedings of the 12th international congress of phonetic sciences, August 19-24, 1991, Aix-en-Provence, France, 2/5*, 138-141.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics, 17*, 257-282.

Port, R. F., & O'Dell, M. L. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics, 13*, 455-471.

Prince, A. S., & Smolensky, P. (in press). *Optimality theory: Constraint interaction in a generative grammar*. Cambridge, MA: MIT Press.

Protopapas, A., Finney, S., & Eimas, P. D. (1995). Effects of syllabification in the perception of spoken English. *Proceedings of Eurospeech 1995, Madrid, 18-21 September 1995*, 2301-04.

Radeau, M., M., J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition, 17/5*, 525-535.

Rayner, K., & Posnansky, C. (1978). Stages of processing in word identification. *Journal of Experimental Psychology: General, 107/1*, 64-80.

Rice, K. D. (1990). Predicting rule domains in the phrasal phonology. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 289-312). Chicago & London: University of Chicago Press.

Rietveld, A. C. M., & Frauenfelder, U. H. (1987). The effect of syllable structure on vowel duration. *Proceedings of the 11th International Congress on Phonetic Science (Tallinn), 4*, 28-31.

Roelofs, A. (1995). *The WEAVER model of word-form encoding in speech production.* (Unpubl. Ms.). MPI for Psycholinguistics, Nijmegen.

Roelofs, A. (1996). Serial order planning of successive morphemes of a word. *Journal of Memory and Language.*

Rubach, J., & Booij, G. (1990). Syllable structure assignment in Polish. *Phonology, 7*, 121-158.

Schane, S. (1972). Natural rules in phonology. In R. P. Stockwell, & R. K. S. Macauley (Eds.), *Linguistic change and linguistic theory* (pp. 199-229). Bloomington: Indiana University Press.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*, 86-102.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition, 12/4*, 315-328.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry, 11*, 563-605.

Selkirk, E. O. (1982). The syllable. In Hulst, H. van der, Smith, N. (Ed.), *The structure of phonological representations. Part II* (pp. 337-383). Dordrecht: Foris.

Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition, 53*, 91-127.

Shattuck-Hufnagel, S. (1979). Speech orders as evidence for a serial order mechanism in sentence production. In W. E. Cooper, & E. C. T. Walker (Eds.), *Sentence processing* (pp. 295-342). Hillsdale, NJ: Lawrence Erlbaum.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109-136). New York: Springer.

Shattuck-Hufnagel, S. (1987). The role of word onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller, & M. Gopnik (Eds.), *Motor and sensory processing in language* (pp. 17-51). Hillsdale, NJ: Lawrence Erlbaum.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition, 42*, 213-259.

Sievers, E. (1881). *Grundzüge der Phonetik.* Leipzig: Breitkopf & Härtel.

Slis, I. H. (1985). *The voiced-voiceless distinction and assimilation of voice in Dutch.* (Unpublished Doctoral Dissertation). Nijmegen: University of Nijmegen.

Slis, I. H. (1986). Assimilation of voice in Dutch as a function of stress, word boundaries, and sex of speaker and listener. *Journal of Phonetics, 14*, 311-236.

Slis, I. H., & van Heugten, M. (1989). Voiced-voiceless distinction in Dutch fricatives. In H. Bennis, & A. Kemmenade van (Eds.), *Linguistics in the Netherlands 1989* (pp. 123-132). Dordrecht: Foris.

Slowiaczek, L. M., & Dinssen, D. A. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics, 13*, 325-341.

Slowiaczek, L. M., & Szymanska, H. J. (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics, 17*, 205-212.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21/3*, 686-698.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language. Vol. 1* (pp. 143-186). London: Lawrence Erlbaum.

Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition, 14*, 17-26.

Stetson, R. H. (1951). *Motor Phonetics.* (2nd ed.). Amsterdam: North-Holland.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences.* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Tousman, S., & Inhoff, A. (1992). Phonology in multisyllabic word recognition. *Journal of Psycholinguistic Research, 21*, 525-544.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15*, 49-74.

Treiman, R. (1984). On the status of final conconant clusters in English syllables. *Journal of verbal Learning and Verbal Behaviour, 23*, 343-356.

Treiman, R., & Danis, C. (1988a). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27*, 87-104.

Treiman, R., & Danis, C. (1988b). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 145-152.

Trommelen, M. (1984). *The syllable in Dutch.* Dordrecht: Foris.

Turk, A. (1994). Articulatory phonetic clues to syllable affiliation: gestural characteristics of bilabial stops. In P. A. Keating (Ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form* (pp. 107-135). Cambridge et al.: Cambridge University Press.

Tweedy, J. R., & Lapinski, R. H. (1981). Facilitating word recognition: Evidence for strategic and automatic factors. *Quarterly Journal of Experimental Psychology, 33A*, 51-59.

Tweedy, J. R., Lapinski, R. H., & Schwaneveldt, R. W. (1977). Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. *Memory & Cognition, 5/1*, 84-89.

Underwood, G. (1976). Semantic interference from unattended printed words. *British Journal of Psychology, 67/3*, 327-338.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In Ferguson, C. A., Menn, L., Stoel-Gammon, C. (Ed.), *Phonological development. Models, research, implications* (pp. 393-422). Timonium, MD: York Press.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*, 311-334.

Wijnen, F., Krikhaar, E., & Os, E. den (1994). The (non)realization of unstressed elements in children's utterances: evidence for a rhythmic constraint. *Journal of Child Language, 21*, 59-83.

Wurzel, W. U. (1970). *Studien zur deutschen Lautstruktur.* Berlin: Akademie-Verlag.

Zwitserlood, P. (1991). Syllabic and morphological factors. *Annual Report of the Max Planck Institute for Psycholinguistics 1991*, 36-37.

Zwitserlood, P., Schriefers, H., Lahiri. A., & Donselaar, W. van (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19/2*, 1-12.

APPENDIX A: MATERIALS AND RESULTS

A1 Experiment 1

A1.1 Pretest

Materials and Results[1]

itemnr	noun (translation)	verb (translation)	RT/ms	STDev	correct (of 50)	chosen[a]
1	fluitje (pipe)	snerpen (bite)	913	285	43	
2	korenveld (cornfield)	wuiven (waft)	935	204	42	e
3	acteur (actor)	schminken (make up)	1049	275	42	
4	tandarts (dentist)	boren (drill)	885	267	50	e
5	dichter (poet)	rijmen (rhyme)	808	135	48	e
6	motor (motor)	ronken (rattle)	836	252	49	e
7	tuinder (gardener)	harken (rake)	1022	327	43	
8	kind (child)	spelen (play)	891	203	48	e
9	camera (camera)	filmen (film)	918	242	47	e
10	koningin (queen)	tronen (throne)	892	262	47	e
11	auto (car)	sturen (drive)	1160	222	42	
12	timmerman (carpenter)	schaven (plane)	1160	394	39	
13	non (nun)	knielen (kneel)	848	236	49	e
14	pijp (pipe)	roken (smoke)	820	241	50	e
15	oog (eye)	tranen (water)	911	231	48	e
16	slang (snake)	wurgen (gulp)	832	178	50	e
17	soep (soup)	koken (cook)	813	204	49	e
18	echo (echo)	galmen (echo)	908	267	49	e
19	zeep (soap)	schuimen (foam)	805	117	50	e
20	schilder (painter)	verven (paint)	1012	225	49	e
21	bezem (broom)	vegen (sweep)	849	212	50	e
22	kouwgom (chewing gum)	kleven (stick)	910	224	45	e
23	kapitein (captain)	seinen (radio)	992	275	44	
24	bloem (flower)	fleuren (bloom)	988	257	45	
25	reiziger (traveller)	boeken (book)	881	212	50	e
26	wind (wind)	gieren (howl)	1196	423	33	
27	boot (boat)	zeilen (sail)	1163	348	38	
28	klimop (ivy)	ranken (creep)	834	197	50	e

[1]Full vowels in open syllables are always long in Dutch and orthographically coded by one letter. Closed syllables can contain long vowels, which are coded by two letters (e.g., *troon* /tro:n/ "throne"), or short vowels, which are coded with one letter (e.g., *rok* /rɔk/ "skirt"). Normally, /i,ɪ/ spell <i> (<ie> codes /i/ in closed syllables), /y,ʏ/ spell <u>, /u/ spells <oe>, /e,ɛ/ spell <e>, /ø/ spells <eu>, /o,ɔ/ spell <o>, /a,ɑ/ spell <a> (the second phoneme is always the short variant).

29	vogel (bird)	piepen (chirp)	1109	357	31	
30	boer (farmer)	ploegen (plough)	1254	367	35	
31	koe (cow)	kalven (calve)	895	228	46	
32	dronkeman (drunkard)	zwieren (stagger)	994	270	49	e
33	water (water)	golven (surf)	1100	296	47	

[a]e=experimental item for main experiment

A1.2 Main Experiment 1

Materials

SET 1

item nr	item-group	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
1	1	kind (child)	spelen (play)	spee	speel	troo	troon
2	2	koningin (queen)	tronen (throne)	troo	troon	spee	speel
				[tro:]	[tro:n]	[spe:]	[spe:l]
3	3	bezem (broom)	vegen (weep)	vee	veeg	koo	kook
4	4	soep (soup)	koken (cook)	koo	kook	vee	veeg
				[ko:]	[ko:k]	[ve:]	[ve:x]
5	5	motor (engine)	ronken (rattle)	ron	ronk	fil	film
				[rɔŋ]	[rɔŋk]	[fɪl]	[fɪlm]

Fillers:

				unrelated IS			
				short	long	short	long
11	1	vrees (fear)	duchten (fear)	ploe	ploek	heu	heur
12	2	voedsel (food)	eten (eat)	heu	heur	ploe	ploek
13	3	oplossing (solution)	zoeken (look for)	hui	huik	kleu	kleup
14	4	bank (bank)	sparen (save)	kleu	kleup	hui	huik
15	5	vijand (enemy)	haten (hate)	prui	pruit	snee	sneef
16	1	handschoen (glove)	boksen (box)	snee	sneef	prui	pruit
17	2	ruzie (squabble)	pesten (torment)	knaa	knaam	soe	soer
18	3	gevaar (danger)	mijden (avoid)	soe	soer	knaa	knaam
19	4	schutter (rifleman)	richten (point)	pru	prun	jaa	jaan
20	5	kleding (clothing)	vouwen (fold)	jaa	jaan	pru	prun
21	1	auto (car)	sturen (drive)	nel	nelf	foe	foen
22	2	doel (target)	raken (hit)	foe	foen	nel	nelf
23	3	rug (back)	duwen (push)	ja	jal	flij	flijk
24	4	ramp (catastrophe)	vluchten (flee)	flij	flijk	ja	jal
25	5	koffie (coffee)	morsen (grind)	hie	hiek	kner	knerk
26	1	boer (farmer)	ploegen (plough)	kner	knerk	hie	hiek

27	2	spoor (trace)	volgen (trail)	braa	braas	di	din
28	3	onderzoek (examination)	testen (examine)	di	din	braa	braas
29	4	tegenspraak (contradiction)	dulden (countenance)	gau	gaut	splei	spleik
30	5	touw (rope)	knellen (squeeze)	splei	spleik	gau	gaut

Practice (noun - verb, IS):
touw (rope) - knellen (squeeze), pink noise; auto (car) - sturen (drive), nelf; winter (winter) - sneeuwen (snow), snee; koffie (coffee) - morsen (grind), hiek; boer (farmer) - ploegen (plough), pink noise; ruzie (squabble) - pesten (torment), knaam

SET 2

item nr	item-group	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
6	1	camera (camera)	filmen (film)	fil	film	rong	ronk
7	2	non (nun)	knielen (kneel)	knie	kniel	traa	traan
8	3	oog (eye)	tranen (water)	tra	traan	knie	kniel
				[tra:]	[tra:n]	[kni:]	[kni:l]
9	4	pijp (pipe)	roken (smoke)	roo	rook	schui	schuim
10	5	zeep (soap)	schuimen (foam)	schui	schuim	roo	rook
				[sxəʏ]	[sxəʏm]	[ro:]	[ro:k]

Fillers:

				unrelated IS			
				short	long	short	long
31	1	hand (hand)	schudden (shake)	prin	prink	koe	koet
32	2	mening (opinion)	vormen (form)	koe	koet	prin	prink
33	3	zon (sun)	warmen (warm)	klaa	klaap	gol	golp
34	4	persoon (person)	bellen (call)	gol	golp	klaa	klaap
35	5	kerst (christmas)	wensen (wish)	roe	roen	ir	irg
36	1	kip (chicken)	braden (fry)	ir	irg	roe	roen
37	2	huid (skin)	krabben (scratch)	sloe	sloet	zi	zik
38	3	koper (copper)	glanzen (shine)	zi	zik	sloe	sloet
39	4	deun (tune)	fluiten (whistle)	luu	luut	bris	brist
40	5	zwemvest(lifejacket)	redden (rescue)	bris	brist	luu	luut
41	1	wasknijper (clothes-peg)	klemmen (hang up)	sui	suin	por	porf
42	2	school (school)	leren (learn)	por	porf	sui	suin
43	3	grap (joke)	lachen (laugh)	zwei	zweil	muu	muun
44	4	tekening (drawing)	schetsen (sketch)	muu	muun	zwei	zweil
45	5	video (video)	huren (borrow)	woe	woen	ai	aig
46	1	wapen (weapon)	moorden (murder)	ai	aig	woe	woen
47	2	butler (butler)	dienen (serve)	zou	zouk	la	las
48	3	schouder (shoulder)	steunen (lean on)	la	las	zou	zouk
49	4	landschap (scenery)	glooien (gleam)	bou	bouk	ul	ulg
50	5	groep (group)	leiden (lead)	ul	ulg	bou	bouk

Practice (noun - verb, IS):
zon (sun) - warmen (warm), pink noise; kip (chicken) - braden (fry), irg; groep (group) - leiden (lead), bou; boot (boat) - rooien (row), roe; video (video) - huren (borrow), pink noise; hand (hand) - schudden (shake), prink

A2 Experiment 2

A2.1 Materials

item number	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
1	tandarts (dentist)	boren (drill)	boo	boor	rij	rijm
2	dichter (poet)	rijmen (rhyme)	rij	rijm	boo	boor
			[rɛi]	[rɛim]	[bo:]	[bo:r]
3	motor (engine)	ronken (rattle)	ron	ronk	fil	film
4	camera (camera)	filmen (film)	fil	film	ron	ronk
			[fɪl]	[fɪlm]	[rɔŋ]	[rɔŋk]
5	pijp (pipe)	roken (smoke)	roo	rook	schui	schium
6	zeep (soap)	schuimen (foam)	schui	schuim	roo	rook
			[sxəʏ]	[sxəʏm]	[ro:]	[ro:k]
7	kind (child)	spelen (play)	spee	speel	troo	troon
8	koningin (queen)	tronen (throne)	troo	troon	spee	speel
			[tro:]	[tro:n]	[spe:]	[spe:l]
9	reiziger (traveller)	boeken (book)	boe	boek	wui	wuiv
10	korenveld (cornfield)	wuiven (wave)	wui	wuiv	boe	boek
			[ʋəʏ]	[ʋəʏv]	[bu:]	[bu:k]
11	klimop (ivy)	ranken (rank)	ran	rank	ver	verv
12	schilder (painter)	verven (pait)	ver	verv	ran	rank
			[vɛr]	[vɛrv]	[rɑŋ]	[rɑŋk]
13	bezem (broom)	vegen (sweep)	vee	veeg	koo	kook
14	soep (soup)	koken (cook)	koo	kook	vee	veeg
			[ko:]	[ko:k]	[ve:]	[ve:x]
15	non (nun)	knielen (kneel)	knie	kniel	traa	traan
16	oog (eye)	tranen (water)	traa	traan	knie	kniel
			[tra:]	[tra:n]	[kni:]	[kni:l]
17	slang (snake)	wurgen (gulp)	wur	wurg	gal	galm
18	echo (echo)	galmen (echo)	gal	galm	wur	wurg
			[xɑl]	[xɑlm]	[ʋør]	[ʋørx]
19	lijm (glue)	kleven (glue)	klee	kleev	zwier	zwier
20	dronkeman (drunkard)	zwieren (stagger)	zwie	zwier	klee	kleef
			[zʋi:]	[zʋi:r]	[kle:]	[kle:v]

Practice (noun- verb, IS):
klok (clock) - tikken (tick), pink noise; hond (dog) - barken (bark), pink noise; fluitje (pipe) - snerpen (smart), sner; timmerman (carpenter) - schaven (plain), kal; koe (cow) - kalven (calve),

kalv; haan (cock) - kraaien (crow), schav

A3 Experiment 4

A 3.1 Pretest

Materials and Results

itemnr	noun (translation)	verb (translation)	RT/ms	STDev	correct chosen (of 120)	
1	mening (oppinion)	achten (respect)	1120	324	71	
2	bedelaar (beggar)	danken (thank)	1053	353	106	e
3	water (water)	drenken (drink)	1110	336	94	e
4	vrees (fear)	duchten (fear)	1148	350	83	e
5	vuilnisbelt (rubbish dump)	dumpen (dump)	899	169	114	e
6	ongeluk (accident)	helpen (help)	1138	294	106	e
7	schutter (rifleman)	richten (point)	1010	289	108	e
8	kan (pitcher)	schenken (pour)	811	224	114	e
9	kwaliteit (quality)	schiften (select)	1114	340	78	
10	ruzie (squabble)	pesten (torment)	1264	349	61	e
11	mes (knife)	slachten (slaughter)	1066	308	100	e
12	vuist (fist)	stompen (punch)	940	285	102	e
13	onderzoek (examination)	testen (examine)	1070	313	96	e
14	verdriet (grief)	troosten (comfort)	1005	288	98	e
15	vinger[2] (finger)	wenken (wave)	1019	295	76	e
16	verjaardag (birhtday)	wensen (wish)	1003	311	107	e
17	detective (detective)	volgen (follow)	912	210	105	e
18	camera (camera)	filmen (film)	807	172	114	e
19	vraag (question)	polsen (enquire)	1182	289	61	
20	hout (wood)	spalken (splint)	1029	268	98	
21	rechter (judge)	schorsen (reverse)	1152	360	91	e
22	oppositie (opposition)	tarten (challenge)	962	196	101	e
23	klei (clay)	vormen (form)	909	235	113	e
24	zon (sub)	warmen (warm)	860	213	115	e

Practice (noun - verb):
touw (rope) - binden (bind), butler (butler) - dienen (serve), vijand (enely) - haten (hate), lineaal (ruler) - meten (measure), pen (pen) - schrijven (write); rekening (bank account) - sparen (save)

[2]After the pretest, a better trigger noun was found (*hand* "hand"). Therefore, this item was included despite its low number of correct responses in the pretest.

A3.2 Materials of Main Experiment 4

SET 1

item number	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
1	bedelaar (beggar)	danken (thank)	dan	dank	vor	vorm
2	klei (clay)	vormen (form)	vor	vorm	dan	dank
			[vɔr]	[vɔrm]	[dɑŋ]	[dɑŋk]
3	ruzie (squabble)	pesten (torment)	pes	pest	rich	richt
4	schutter (rifleman)	richten (point)	rich	richt	pes	pest
			[rɪç]	[rɪçt]	[pɛs]	[pɛst]

Fillers:

			unrelated IS			
			short	long	short	long
21	bezem (broom)	vegen (sweep)	ploe	ploek	gau	gaut
22	timmerman (carpenter)	schaven (plain)	prui	pruit	kie	kieg
23	tandarts (dentist)	boren (drill)	kie	kieg	nel	nelf
24	reiziger (traveller)	boeken (book)	slie	sliet	tui	tuik
25	wasknijper (clothes-peg)	hangen (hang up)	ir	irg	flie	flien
26	belofte (promose)	zweren (swear)	gol	golp	ir	irg
27	huurder (renter)	huren (rent)	nel	nelf	gla	glam
28	spijskaart (menu)	kiezen (choose)	pru	prun	heu	heur
29	brievenbus (letter box)	legen (put)	pu	pur	lar	larp
30	pen (pen)	schrijven (write)	heu	heur	prui	pruit
31	kruispunt (crossing)	stoppen (stop)	tui	tuik	di	din
32	touw (rope)	binden (bind)	stu	stul	kie	kieg
33	poederdoos (powder compact)	schminken (make up)	lar	larp	stu	stul
34	neus (nose)	ruiken (smell)	gla	glam	tui	tuik

SET 2

item number	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
5	verdriet (grief)	troosten (comfort)	troo	troost	slach	slacht
6	mes (knife)	slachten (slaughter)	slach	slacht	troo	troost
			[slɑx]	[slɑxt]	[tro:]	[tro:st]
7	verjaardag (birthday)	wensen (wish)	wen	wens	dum	dump
8	vuilnisbelt (rubbish dump)	dumpen (dump)	dum	dump	wen	wens
			[døm]	[dømp]	[ʋɛn]	[ʋɛns]

Fillers:

				unrelated IS		
			short	long	short	long
35	lijm (glue)	kleven (glue)	flie	flien	knaa	knaam
36	telefoon (phone)	bellen (ring)	di	din	grie	griem
37	klauw (claw)	grijpen (gralp)	knaa	knaam	hie	hiek
38	stem (voice)	roepen (call)	tui	tuik	grie	griem
39	pap (porridge)	voeden (feed)	hie	hiek	ja	jal
40	vraag (question)	vragen (ask)	grie	griem	nau	naum
41	bank (bank)	sparen (save)	foe	foen	flij	flijk
42	kruier (porter)	dragen (carry)	gee	geep	ai	aig
43	rug (back)	duwen (push)	ja	jal	gee	geep
44	krans (wreathe)	eren (honor)	prin	prink	dro	drok
45	olie (oil)	zalven (anoint)	vu	vum	trei	treik
46	portier (doorkeeper)	melden (report)	flij	flijk	zol	zolm
47	speurtocht (search)	vinden (find)	kner	knerk	prin	prink
48	vinger (finger)	wijzen (point)	gau	gaut	ploe	ploek

SET 3

item number	noun (translation)	verb (translation)	related IS short	long	unrelated IS short	long
9	zon (sub)	warmen (warm)	war	warm	hel	help
10	ongeluk (accident)	helpen (help)	hel	help	war	warm
			[hɛl]	[hɛlp]	[ʋɑr]	[ʋɑrm]
11	detective (detective)	volgen (follow)	vol	volg	tar	tart
12	oppositie (opposition)	tarten (challenge)	tar	tart	vol	volg
			[tɑr]	[tɑrt]	[vɔl]	[vɔlx]

Fillers:

				unrelated IS		
			short	long	short	long
49	schop (schovel)	graven (dig)	dro	drok	klaa	klaap
50	bezem (broom)	vegen (sweep)	jon	jons	fraa	fraas
51	vlees (meat)	keuren (test)	trei	treik	braa	braas
52	koper (buyer)	kopen (buy)	soe	soer	nir	nirk
53	personeel (personal)	werven (recruit)	zol	zolm	ul	ulg
54	bakker (baker)	bakken (bake)	sloe	sloet	zi	zik
55	pastoor (priest)	dopen (christen)	klaa	klaap	soe	soer
56	geweld (force)	breken (break)	fraa	fraas	il	ilf
57	schaar (scissors)	knippen (cut)	nau	naum	lu	lur
58	school (school)	leren (learn)	braa	braas	sloe	sloet
59	etalage (shop-window)	tonen (show)	il	ilf	kleu	kleup
60	waag (balance)	wegen (weigh)	nir	nirk	splei	splei
61	mond (mouth)	zoenen (kiss)	luu	luut	fui	fuis

62	tang (tongs)	knijpen (pinch)	bou	bouk	far	farn

SET 4

item number	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
13	camera (camera)	filmen (film)	fil	film	schen	schenk
14	kan (pitcher)	schenken (pour)	schen	schenk	fil	film
			[sxɛŋ]	[sxɛŋk]	[fɪl]	[fɪlm]
15	onderzoek (examination)	testen (examine)	tes	test	stom	stomp
16	vuist (fist)	stompen (punch)	stom	stomp	tes	test
			[stɔm]	[stɔmp]	[tɛs]	[tɛst]

Fillers:

			unrelated IS short	unrelated IS long	unrelated IS short	unrelated IS long
63	kind (child)	spelen (play)	ai	aig	foo	fook
64	belediging (insult)	krenken (offend)	lu	lur	stoo	stoop
65	koffer (suitcase)	pakken (pack)	kleu	kleup	dil	dils
66	verhaal (story)	boeien (captivate)	ul	ulg	mui	muik
67	bode (messenger)	brengen (bring)	por	porf	hui	huik
68	gebod (order)	dwingen (force)	splei	spleik	ruu	ruuf
69	oor (ear)	horen (hear)	trau	traut	woe	woen
70	schilder (painter)	malen (paint)	woe	woen	bou	bouk
71	haak (hook)	trekken (pull)	zi	zik	spoo	spool
72	caroussel (merry-go-round)	draaien (turn)	foo	fook	hui	huik
73	vriend (friend)	kennen (know)	hui	huik	woe	woen
74	hout (wood)	spalken (splint)	fui	fuis	li	lijs
75	gevaar (danger)	mijden (avoid)	raa	raag	zwei	zweil
76	strijkijzer (iron)	strijken (iron)	far	farn	maa	maap

SET 5

item number	noun (translation)	verb (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
17	hand (hand)	wenken (wave)	wen	wenk	duch	ducht
18	vrees (fear)	duchten (fear)	duch	ducht	weng	wenk
			[døx]	[døxt]	[ʋɛŋ]	[ʋɛŋk]
19	water (water)	drenken (water)	dren	drenk	schor	schors
20	rechter (judge)	schorsen (reverse)	schor	schors	dreng	drenk
			[sxɔr]	[sxɔrs]	[drɛŋ]	[drɛŋk]

Fillers:

				unrelated IS		
			short	long	short	long
77	kathedraal (cathedral)	prijzen (celebrate)	stoo	stoop	praa	praak
78	liniaal (ruler)	meten (measure)	sui	suin	raa	raag
79	eten (meal)	proeven (prove)	bou	bouk	far	farn
80	butler (butler)	dienen (serve)	spoo	spool	muu	muun
81	nagel (nail)	krabben (scratch)	zwei	zweil	sui	suin
82	oven (oven)	braden (roast)	bris	brist	mui	muik
83	zwemvest (life-jacket)	redden (rescue)	mui	muik	maa	maap
84	vel (skin)	strelen (fondle)	maa	maap	rouu	ruuf
85	kraan (crane)	tillen (lift)	ruu	ruuf	praa	praak
86	postzegel (stamp)	zenden (send)	praa	praak	baa	baap
87	mening (oppinion)	uiten (utter)	baa	baap	muu	muun
88	knie (knee)	smeken (beg)	muu	muun	lij	lijs
89	som (sum)	snappen (get)	lij	lijs	dil	dils
90	lap (cloth)	poetsen (clean)	dil	dils	bou	bouk

A3.3 Mean RTs per set in Experiment 4

	SOA					
	0 ms		150 ms		300 ms	
	IS length					
IS	short	long	short	long	short	long
	SET1					
related	825	848	802	837	856	906
unrelated	934	955	977	974	944	875
Δ	109	107	175	137	88	-31
	SET 2					
related	811	835	805	795	866	872
unrelated	966	961	918	937	932	941
Δ	155	126	113	142	66	69
	SET 3					
related	819	847	796	829	861	928
unrelated	913	927	924	977	926	987
Δ	94	80	128	148	65	59
	SET 4					
related	824	831	753	767	826	848
unrelated	863	870	880	848	864	856
Δ	39	39	127	81	38	-8

Note. The values represent ms, Δ symbolizes the mean difference scores (unrelated - related).

A4 Experiment 5

A4.1 Pretest

Materials and Results

itemnr	set	noun (translation)	verb (translation)	RT/ms	STDev	correct (of 40)	chosen
1	1	glas (glass)	tinten (tint)	823	243	40	e
2	1	dekbed (quilt)	luchten (air)	715	110	40	e
3	1	gat (hole)	boren (drill)	795	205	37	e
4	1	paasei (easter-egg)	zoeken (search)	691	106	38	e
5	1	bloempje (plant)	planten (plant)	842	221	36	
6	1	kolbertje (jacket)	lenen (borrow)	747	154	39	e
7	1	vakantiehuis (vacation home)	boeken (book)	794	197	40	e
8	1	gedicht (poem)	leren (learn)	804	230	39	e
9	1	autootje (car)	keuren (test)	865	246	36	
10	1	succes (success)	vieren (celebrate)	769	103	38	e
11	1	pijpje (pipe)	roken (smoke)	715	113	40	e
12	1	bier (beer)	koelen (cool)	870	261	38	
13	1	bedrag (amout)	noemen (name)	718	119	38	e
14	1	hout (wood)	lijmen (glue)	752	172	40	e
15	1	racisme (racism)	haten (hate)	690	96	40	e
16	1	beslag (batter)	roeren (stir)	780	184	35	
17	2	grind (gravel)	harken (rake)	715	143	40	
18	2	gas (gas)	pompen (pump)	780	129	39	e
19	2	kind (child)	dopen (christen)	853	205	40	e
20	2	paard (horse)	tomen (bridle)	933	260	35	
21	2	kado (present)	ruilen (exchange)	938	310	36	
22	2	eten (meal)	koken (cook)	796	196	39	e
23	2	vee (cattle)	slachten (slaughter)	779	232	38	
24	2	wondje (wound)	hechten (suture)	760	135	39	e
25	2	ongeluk (accident)	duchten (fear)	992	342	39	
26	2	touw (rope)	knopen (knot)	792	197	39	e
27	2	plaatje (picture)	kleuren (color)	829	203	38	e
28	2	raam (window)	sluiten (close)	668	79	40	e
29	2	liedje (song)	fluiten (whistle)	870	174	37	
30	2	korset (corset)	snoeren (lace up)	776	195	36	e
31	2	vest (waistcoat)	haken (crochet)	926	259	31	
32	2	gewicht (weight)	meten (measure)	757	121	40	e

A4.2 Materials of Main Experiment 5

			related IS		unrelated IS	
itemnr	verb (translation)	noun (translation)	short	long	short	long
1	glas (glass)	tinten (tint)	tin	tint	luch	lucht
2	dekbed (quilt)	luchten (air)	luch	lucht	tin	tint
			[løx]	[løxt]	[tɪn]	[tɪnt]
3	water (water)	pompen (pump)	pom	pomp	hech	hecht
4	wondje (wound)	hechten (suture)	hech	hecht	pom	pomp
			[hɛx]	[hɛxt]	[pɔm]	[pɔmp]
5	gat (hole)	boren (drill)	boo	boor	zoe	zoek
6	paasei (easter-egg)	zoeken (search)	zoe	zoek	boo	boor
			[zu:]	[zu:k]	[bo:]	[bo:r]
7	vakantiehuis (vacation home)	boeken (book)	boe	boek	lee	leer
8	gedicht (poem)	leren (learn)	lee	leer	boe	boek
			[le:]	[le:r]	[bu:]	[bu:k]
9	succes (success)	vieren (celebrate)	vie	vier	noe	noem
10	bedrag (amount)	noemen (name)	noe	noem	vie	vier
			[nu:]	[nu:m]	[vi:]	[vi:r]
11	pijpje (pipe)	roken (smoke)	roo	rook	lij	lijm
12	hout (wood)	lijmen (glue)	lij	lijm	roo	rook
			[lɛi]	[lɛim]	[ro:]	[ro:k]
13	eten (meal)	koken (cook)	koo	kook	mee	meet
14	gewicht (weight)	meten (measure)	mee	meet	koo	kook
			[me:]	[me:t]	[ko:]	[ko:k]
15	kolbertje (jacket)	lenen (borrow)	lee	leen	doo	doop
16	kind (child)	dopen (christen)	doo	doop	lee	leen
			[do:]	[do:p]	[le:]	[le:n]
17	touw (rope)	knopen (knot)	knoo	knoop	slui	sluit
18	raam (window)	sluiten (close)	slui	sluit	knoo	knoop
			[slɵʏ]	[slɵʏt]	[kno:]	[kno:p]
19	plaatje (picture)	kleuren (color)	kleu	kleur	snoe	snoer
20	korset (corset)	snoeren (lace up)	snoe	snoer	kleu	kleur
			[snu:]	[snu:r]	[klø]	[klør]

Practice (noun- verb, IS):
vee (cattle) - slachten (slaughter), pink noise; racisme (racism) - haten (hate), pink noise; beslag (batter) - roeren (stir), roe; bloempje (flower) - planten (plant), plant; kado (present) - ruilen (exchange), flui; liedje (song) - fluiten (whistle), ruil

A5 Experiment 6

A5.1 Pretest: Materials

SET 1

cue - morphologically simple target:
kam (comb) - borstel (brush), danseres (dancer) - gratie (grace), baron (baron) - hertog (duke), burcht (castle) - kerker (dungeon), getjilp (chirping) - krekel (grasshopper), raaf (raven) - merel (blackbird), rekening (account) - nota (bill), schelp (shell) - parel (pearl), klepper (rattle) - ratel (rattle), eten (meal) - schotel (dish), bit (bit) - teugel (rein)

cue - morphologically complex target:
zeilen (sails) - boten (boats), struiken (shrubs) - hagen (hedge), vliegen (flies) - horren (screens), barakken (sheds) - keten (huts), fotos (photographs) - lijsten (frames), hooi (hay) - mijten (ricks), omheiningen (fences) - palen (poles), bogen (bows) - pijlen (arrows), nachten (nights) - schimmen (shadows), voeten (feet) - sokken (socks), wespen (wasps) - steken (stings), tranen (tears) - uien (onions)

SET 2

cue - morphologically complex target:
vrouwen (women) - borsten (breasts), vissen (fish) - graten (fishbones), geweien (antlers) -herten (deer), pastoren (priests) - kerken (churches), geruis (murmur) - kreken (stream), ponten (ferries) - meren (seas), eekhoorns (squirrels) - noten (nuts). bruiloften (weddings) - paren (pairs), honing (honey) - raten (honeycombs), geweren (guns) - schoten (shots), slokken (swallows) - teugen (draughts)

cue - morphologically simple target:
vet (fat) - boter (butter), sneeuw (snow) - hagel (hale), angst (fear) - horror (horror), fornuis (cooker) - ketel (kettle), lijster (thrush) - nest (nest), paus (pope) - mijter (mitre), viskraampje (fish booth) - paling (eel), brug (bridge) - pijler (pole), yoghurt (yoghurt) - schimmel (mold), beeld (statue) - sokkel (plinth), cactus (cactus) - stekel (prickle), koe (cow) - uier (udder)

Practice (cue - target):
papieren (papers) - pennen (pencils), stad (town) - straat (street), gebak (pastry) - cake (cake)

A5.2 Pretest: Results

pair	itnr	target	err	rt	itnr	target	err	rt	diffrt
> 1	1	borstel	0	750	24	borsten	0	756	- 6
> 2	2	gratie	0	786	25	graten	4	735	51
3	3	hertog	0	823	26	herten	1	949	126*
> 4	4	kerker	1	991	27	kerken	1	951	40

5	5	krekel	0	846	28	kreken	-	(963 sg)	(-117*)
6	6	merel	0	897	29	meren	4	**<u>1073</u>**	-176*
>7	7	nota	0	792	30	noten	0	858	- 66
>8	8	parel	2	898	31	paren	1	800	98
>9	9	ratel	0	842	32	raten	-	(820 sg)	(22)
>10	10	schotel	1	838	33	schoten	0	829	9
>11	11	teugel	0	760	34	teugen	2	853	- 93
>12	35	boter	2	737	12	boten	0	784	- 47
13	36	hagel	3	756	13	hagen	4	910	-154*
14	37	horror	2	774	14	horren	**<u>7</u>**	933	-159*
15	38	ketel	0	788	15	keten	3	945	-157*
16	(39	lijster			16	lijsten)			
17	40	mijter	2	**<u>1002</u>**	17	mijten	1	762	240*
>18	41	paling	1	912	18	palen	2	875	37
19	42	pijler	2	949	19	pijlen	0	794	155*
>20	43	schimmel	0	719	20	schimmen	1	780	- 61
>21	44	sokkel	0	792	21	sokken	0	708	84
>22	45	stekel	1	777	22	steken	0	747	30
>23	46	uier	0	773	23	uien	2	851	- 78

Bold underlined type indicates the RTs that lead to the exclusion of the target pair from the materials of the main experiment. A '*' marks the pairs where the difference in RT for the complex and the simple member exceeds 100 ms. Furthermore, in item 16 cue and target had been exchanged in the pretest by accident. Nevertheless the pair was included in the materials later for the main experiment. A '>' in the beginning of the line marks the 16 'good pairs' that was included in the analysis of the main experiment.

A5.3 Materials of Main Experiment 6

Morphologically Complex Targets:

itemnr	cue (translation)	target (translation)	related IS short	related IS long	unrelated IS short	unrelated IS long
1	eekhoorn (squirrel)	noten (nuts)	noo	noot	paa	paal
11	omheining (fence)	palen (poles)	paa	paal	noo	noot
			[pa:]	[pa:l]	[no:]	[no:t]
2	honing (honey)	raten (honeycombs)	raa	raat	teu	teug
12	slok (swallow)	teugen (draughts)	teu	teug	raa	raat
			[tø:]	[tøx]	[ra:]	[ra:t]
3	kapitein (captain)	boten (boats)	boo	boot	paa	paar
13	bruiloft (wedding)	paren (pairs)	paa	paar	boo	boot
			[pa:]	[pa:r]	[bo:]	[bo:t]

4	barak (shed)	keten (huts)	kee	keet	lij	lijst
14	foto (photograph)	lijsten (frames)	lij	lijst	kee	keet
			[lɛɪ]	[lɛɪst]	[ke:]	[ke:t]
5	wesp (wasp)	steken (stings)	stee	steek	graa	graat
15	vis (fish)	graten (bones)	graa	graat	stee	steek
			[xra:]	[xra:t]	[ste:]	[ste:k]
6	geruis (murmur)	kreken (streams)	kree	kreek	schoo	schoot
16	baby (baby)	schoten (laps)	schoo	schoot	kree	kreek
			[sxo:]	[sxo:t]	[kre:]	[kre:k]
7	pastoor (priest)	kerken (churches)	ker	kerk	bor	borst
17	vrouw (woman)	borsten (breasts)	bor	borst	ker	kerk
			[bɔr]	[bɔrst]	[kɛr]	[kɛrk]
8	traan (tear)	uien (onions)	u	ui	so	sok
18	voet (foot)	sokken (socks)	so	sok	u	ui
			[sɔ]	[sɔk]	[ə]	[əʏ]
9	nacht (night)	schimmen (shadows)	schi	schim	her	hert
19	gewei (antler)	herten (deer)	her	hert	schi	schim
			[hɛr]	[hɛrt]	[sxɪ]	[sxɪm
10	struik (shrub)	hagen (hedges)	haa	haag	pij	pijl
20	boog (bow)	pijlen (arrows)	pij	pijl	haa	haag
			[pɛɪ]	[pɛɪl]	[ha:]	[ha:x]

Morphologically Simple Targets:

21	rekening (account)	nota (bill)	noo	noot	paa	paal
31	viskraampje (fosh booth)	paling (eel)	paa	paal	noo	noot
			[pa:]	[pa:l]	[no:]	[no:t]
22	klepper (rattle)	ratel (rattle)	raa	raat	teu	teug
32	slok (swallow)	teugen (draughts)	teu	teug	raa	raat
			[tø]	[tøx]	[ra:]	[ra:t]
23	vet (fat)	boter (butter)	boo	boot	paa	paar
33	schelp (shell)	parel (pearl)	paa	paar	boo	boot
			[pa:]	[pa:r]	[bo:]	[bo:t]
24	fornuis (cooker)	ketel (kettle)	kee	keet	lij	lijst
34	nest (nest)	lijster (thrush)	lij	lijst	kee	keet
			[lɛɪ]	[lɛɪst]	[ke:]	[ke:t]
25	cactus (cactus)	stekel (prickle)	stee	steek	graa	graat
35	danseres (danser)	gratie (grace)	graa	graat	stee	steek
			[xra:]	[xra:t]	[ste:]	[ste:k]
26	getjilp (murmur)	krekel (stream)	kree	kreek	schoo	schoot
36	eten (meal)	schotel (dish)	schoo	schoot	kree	kreek
			[sxo:]	[sxo:t]	[kre:]	[kre:k]
27	burcht (castle)	kerker (dungeon)	ker	kerk	bor	borst
37	kam (comb)	borstel (brush)	bor	borst	ker	kerk
			[bɔr]	[bɔrst]	[kɛr]	[kɛrk]

28	koe (cow)	uier (udder)	u	ui	so	sok
38	beeld (statue)	sokkel (plinth)	so	sok	u	ui
			[sɔ]	[sɔk]	[ə]	[əʏ]
29	yoghurt (yoghurt)	schimmel (mold)	schi	schim	her	hert
39	baron (baron)	hertog (duke)	her	hert	schi	schim
			[hɛr]	[hɛrt]	[sxɪ]	[sxɪm]
30	sneeuw (snow)	hagel (hale)	haa	haag	pij	pijl
40	brug (brug)	pijler (pillar)	pij	pijl	haa	haag
			[pɛɪ]	[pɛɪl]	[ha:]	[ha:x]

Practice (cue - target, IS):
auto (car) - banden (tires), ban; vuilnis (rubbish) - bezems (brooms), beez; lied (song) - strofen (strophes), stroo; lied (song) - strofen (strophes), ban; hond (dog) - liijnen (line), beez

A6 Experiment 7

A6.1 Materials

item	noun (transl.) - verb (transl.)	related IS word	related IS pseudoword	unrelated IS word	unrelated IS pseudoword
1	gat (hole) - boren (drill)	boom	book	meel	meef
2	gewicht (weight) - meten (measure)	meel (flour)	meef	boom (tree)	book
		[me:l]	[me:f]	[bo:m]	[bo:k]
3	vakantiehuis (holiday home) - boeken (book)	boer	boep	leed	leel
4	kolbertje (jacket) - lenen (borrow)	leed (sorrow)	leel	boer (farmer)	boep
		[le:t]	[le:l]	[bu:r]	[bu:p]
5	gedicht (poem) - leren (learn)	leek	leel	noen	noet
6	bedrag (amount) - noemen (name)	noen (noon)	noet	leek (layman)	leel
		[nu:n]	[nu:t]	[le:k]	[le:l]
7	pijpje (pipe) - roken (smoke)	roos	roog	lijf	lijg
8	hout (wood) - lijmen (glue)	lijf (body)	lijg	roos (rose)	roog
		[lɛɪ]	[lɛɪx]	[ro:s]	[ro:x]
9	eten (meal) - koken (cook)	koor	koom	vies	vien
10	succes (success) - vieren (celebrate)	vies (dirty)	vien	koor (choire)	koom
		[vi:s]	[vi:n]	[ko:r]	[ko:m]
11	kind (child) - dopen (christen)	doos	doog	zoen	zoel
12	paasei (easter egg) - zoeken (search)	zoen	zoel	doos	doog
		[zu:]	[zu:l]	[do:s]	[do:x]
13	touw (rope) - knopen (knot)	knook	knoot	sluis	sluin
14	raam (window) - sluiten (close)	sluis (lock)	sluin	knook (bone)	knoot
		[sləʏs]	[sləʏn]	[kno:k]	[kno:t]
15	korset (corset) - snoeren (lace up)	snoek	snoen	kleum	kleus
16	plaatje (picture) - kleuren (color)	kleum (shiver)	kleus	snoek (pike)	snoen
		[kløm]	[kløs]	[snu:k]	[snu:n]

17	onweer (thunderstorm) - haten (hate)	haar	haap	ruit	ruip
18	kado (present) - ruilen (exchange)	ruit	ruip	haar	haap
		[rəYt]	[rəYp]	[ha:r]	[ha:p]

Practice (noun - verb, IS):
beslag (dough) - roeren (stir), roep (call), roen; bloempje (flower) - planten (plant), plank (plank), plamp; glas (glass) - tinten (tint), tink (tinkle)

A7.1 Table of Experiments 1 to 7

(see following page)
Notes:
(a) interaction verbform by relatedness by length: significant at .05-level at SOA 0
(b) no length variation

A7.2 Difference Scores from Pink Noise Baseline

		target verb form			
		infinitive		past tense	
		IS length			
SOA	relatedness	short	long	short	long
		Experiment 1			
0 ms	related	58	56	35	10
	unrelated	- 7	- 3	-19	-32
150 ms	related	19	24	38	22
	unrelated	-73	-64	-57	-61
		Experiment 2			
0 ms	related	85	67	58	61
	unrelated	-52	-71	-59	-100
150 ms	related	53	47	43	45
	unrelated	-61	-81	-79	- 95
		Experiment 3			
-150 ms	related	78	67	65	58
	unrelated	-47	-76	-63	- 80
0 ms	related	79	83	93	63
	unrelated	-72	-91	-74	- 97
150 ms	related	99	86	93	78
	unrelated	-49	-66	-74	-113
300 ms	related	64	52	60	56
	unrelated	-59	-57	-37	- 38

Note. Values represent mean difference scores in ms. The experiments tested verb forms like *koken* ("to cook") and *kookte* ("cooked").

Experiment, *targets*	main effect relatedness	main effect length	mean RT in ms	mean length of IS	% related IS	SOAs	number of targets	interaction relatedness by length	comments
1, *koken, kookte*	all F **	n.s.	807	455	10	0, 150	10	n.s.	
2, *koken, kookte*	all F **	all F **	814	450	40	0, 150	20	all F *	(a)
3, *koken, kookte*	all F **	all F **	786	310	40	-150, 0, 150, 300	20	n.s.	
4, *kook 't*	all F **	F_1 **, F_2 n.s.	877	440	12.5	0, 150, 300	16	n.s.	
5a, *kook 't*	all F **	all F **	849	510	40	-150, 0, 150	20	n.s.	
5b, *kookt 't*	all F **	n.s.	844	510	40	-150, 0	12	n.s.	
6, *boten, boter*	all F **	n.s.	800	422	40	0, 150, 300	14	all F **	morphology control
7, *koken*	all F **	(b)	821	560	40	0, 150	18	n.s.	lexical control

		Experiment					
		5a		5a (12 items)		5b	
		IS length					
SOA	IS	short	long	short	long	short	long
-150 ms	related	71	29	45	27	28	26
	unrelated	-41	-72	-34	-68	-77	-46
0 ms	related	72	49	64	53	42	41
	unrelated	-57	-92	-45	-77	-75	-79
150 ms	related	34	39	17	25	---	---
	unrelated	-70	-53	-51	-35	---	---

Note. Values represent mean difference scores in ms. The experiments tested targets like *kook het* ("cook it") and *kookt het* ("cook (pl.) it").

	Experiment 6		
		IS length	
SOA	IS	short	long
0 ms	related	37	39
	unrelated	-45	-61
150 ms	related	41	51
	unrelated	-48	-64
300 ms	related	7	15
	unrelated	-64	-68

*Note.*The experiment tested words like *boten* ("boats") and *boter* ("butter"). The values represent mean difference scores in ms over complex and simple words.

A8 Experiments 8a,b

A8.1 Materials of the First Production Experiment (8a)

List of all minimal pairs, incuding information on genus (m = masculine, f = feminine, n = neuter), plural affix (--- = no plural), mass noun (y = yes, - = no), frequency

pair	noun (translation)	gen	plur	mass	freq/1 mill
1	bod (offer)	n	---	-	9
	bot (bone)	n	-en	-	19
2	bond (bond)	m	-en	-	20
	bont (fur)	n	---	y	2
3	ford (Ford)	m	-s	-	10
	fort (fort)	n	-en	-	16
4	graad (degree)	m	-en	-	31
	graat (fishbone)	f	-en	-	2

5	kruid (herb)	n	-en	-	22
	kruit (gunpowder)	n	---	y	2
6	lood (lead)	n	---	y	10
	loot (shoot)	f	-en	-	1
7	nood (poverty)	f	-en/---	-/y	31
	noot (nut)	f	-en	-	19
8	pond (pound)	n	-en	-	15
	pont (ferry)	f	-en	-	4
9	rad (wheel)	n	-eren	-	9
	rat (rat)	f	-en	-	23
10	wand (wall)	m	-en	-	44
	want (mitten)	f	-en	-	1
11	lab (lab)	n	-s	-	3
	lap (cloth)	m	-en	-	13
12	slib (silt)	n	---	y	1
	slip (corner)	f	-en	-	4

Fillers (noun, freq/1 mill - noun, freq/1 mill):
onset
duif (pigeon), 19 - fuif (feasting), 2; troon (throne), 12 - boon (bean), 15; kruik (jug), 7 - luik (hatch), 12; vin (fin), 6 - kin (chin), 32
nucleus
bijl (axe), 11- bel (bell), 34; kas (cash register), 12 - kies (molar), 9; hok (shed), 15 - hak (heel), 12; rem (break), 7 - riem (belt), 22
coda
halm (stalk), 2 - harp (harp), 2; mop (brick), 7 - mot (moth), 7; bek (mouth), 21 - bes (berry), 7; schim (shadow), 12 - schil (bowl), 6

Combinations of first and second noun (here given in the clitic condition, the same combinations were used for the final and the nasal condition) are given below. All sentences were preceded either by *Pien zegt* ("Pien says") or by *Toon zegt* ("Toon says"):

Ik zie een ford en een aapje (I see a Ford and a little monkey), Ik zie een fort en een ader (... a fort and a vein), Ik zie een wand en een ever (... a wall and a boar), Ik zie een want en een ezel (... a mitten and a donkey),
Ik zag een lab en een otter (I saw a lab and an otter), Ik zag een lap en een orgel (... a towel and an organ), Ik zag slib en een egel (... silt and a hedgehog), Ik zag een slip en een eikel (...a corner and a corn),
Er is een bod en een oven (There is an offer and an oven), Er is een bot en een ober (... a bone and a waiter), Er is een graad en een anjer (... a degree and a carnation), Er is een graat en een angel (... a fishbone and a sting),
Er was een bond en een emmer (There was a bond and a basket), Er was bont en een engel (... fur and an angel), Er was nood en een aster (... poverty and an aster), Er was een noot en een arend (... a nut and an eagle),
Ik heb lood en een enkel (I have lead and an ankle), Ik heb een loot en een erker (... a shoot and

a bay), Ik heb een rad en een anker (... a wheel and an anchor), Ik heb een rat en een adder (... a rat and a viper),
Ik had een kruid en een oksel (I had a herb and an armpit), Ik had kruit en een ordner (gunpowder and a file), Ik had een pond en een akker (... a pound and a field), Ik had een pont en een akte (... a ferry and a deed)

Example for the other context conditions:
nasal: Ik had een kruid naast een oksel (I had a herb near an armpit)
final: Ik had een oksel en een kruid (I had an armpit and a herb)

Fillers:

Ik zie een troon en een arts (I see a throne and a doctor), Ik zie een boon en een bagger (... a bean and a dredge), Ik zie een schim en een hoorn (... a shadow and a horn), Ik zie een schil en een beker (I see a bowl and a mug),
Ik zag een kas en een eik (I saw a cash register and an oak), Ik zag een kies en een degen (... a molar and a sword), Ik zag een halm en een helm (...a stalk and a helmet), Ik zag een harp en een deken (... a harp and a blanket),
Er is een hok en een oom (There is a shed and a uncle), Er is een hak en een bever (... a heel and a beaver), Er is een duif en een klok (... a pigeon and a clock), Er is een fuif en een bezem (... a feasting and a broom,
Er was een vin en een oor (There was a fin and an ear), Er was een kin en een deksel (... a chin and a lid), Er was een bek en een hark (... a mouth and a rake), Er was een bes en een fakkel (... a berry and a torch),
Ik heb een kruik en een ui (I have a jug and an onion), Ik heb een luik en een gordel (... a hatch and a belt), Ik heb een bijl en een boon (... an axe and a bean), Ik heb een bel en een gondel (... a bell and a gondola),
Ik had een rem en een uil (I had a break and an owl), Ik had een riem en een lepel (... a belt and a spoon), Ik had een mop en een bank (... a brick and a bank), Ik had een mot en een ladder (... a moth and a ladder

A8.2 Materials of the First Perception Experiment (8b)

The following seven minimal pairs were investigated in the perception task:
bod/bot, ford/fort, graad/graat, lab/lap, pond/pont, rad/rat, wand/want

A8.3 Results for the Voiced Responses ('d' = (4)+(5))

Analysis with 4 speakers:

Voiced responses in the three context conditions:

context	freq.	prop.
clitic	394	.36
final	376	.34
nasal:	689	.62

Effect of context: $F(2;38) = 40.81$, $MSE = .44$, $p < .01$
In all response categories, clitic and final context patterned similarly and differed from the results in the nasal context as was tested in Tukey tests.

Effect of voicing: $F(1,19) = 4.59$, $MSE = .14$, $p < .05$

Reanalysis with 3 speakers:

Effect of context: $F(2;38) = 49.22$, $MSE = .36$, $p < .01$
Effect of voicing: $F(1;19) = 5.58$, $MSE = .13$, $p < .03$

A8.4 Reanalysis of Responses to Items Judged as Encliticized

	response						
voicing	(1) t	(2) t?	(1)+(2)	(3) ?	(4) d?	(5) d	(4)+(5)
				clitic			
voiced	48	62	110	22	28	20	48
	(.27)	(.34)	(.61)	(.12)	(.16)	(.11)	(.27)
voiceless	51	41	92	23	31	34	65
	(.28)	(.23)	(.51)	(.13)	(.17)	(.19)	(.36)
				final			
voiced	64	42	106	13	30	31	61
	(.36)	(.23)	(.59)	(.07)	(.17)	(.17)	(.34)
voiceless	57	40	97	19	38	26	64
	(.32)	(.22)	(.54)	(.11)	(.21)	(.14)	(.36)
				nasal			
voiced	29	21	50	29	57	44	101
	(.16)	(.12)	(.28)	(.16)	(.32)	(.24)	(.56)
voiceless	27	34	61	28	56	35	91
	(.15)	(.19)	(.34)	(.16)	(.31)	(.19)	(.51)

Frequency and Proportion (in Parentheses) for the Different Response Categories in the Three Context and Two Voicing Conditions, for those Minimal Pairs of which both Members had been Judged as Encliticized in the Clitic Context.

A9 Experiments 9a,b

A9.1 Materials of the Second Production Experiment (9a)

itemnr	voiced	voiceless
1	bod en haas (offer and fillet)	bot en haan (bone and cock)
2	bond en eer (bond and honor)	bont en erf (fur and lot)
3	ford en oom (Ford and uncle)	fort en oor (fort and ear)
4	graad en hol (degree and cave)	graat en hof (fishbone and court)

5	kruid en hal (herb and hall)	kruit en hak (gunpowder and heel)
6	lood en ei (lead and egg)	loot en eis (shoot and demand)
7	nood en aak (disease and barge)	noot en aas (nut and prey)
8	pond en os (pound and ox)	pont en olm (ferry and elm)
9	raad en hen (commission and hen)	raat en helm (honeycomb and helmet)
10	rad en ijs (wheel and ice)	rat en ijk (rat and stamp)
11	wand en aar (wall and are)	want en aal (mitten and eal)
12	lab en hoen (lab and chicken)	lap en hoef (cloth and hoof)
13	slib en ui (silt and onion)	slip en uil (corner and owl)

The examples are in the clitic context. The two nouns were exchanged for the final context condition (e.g., *ijk en rat*). Four pairs (graad/graat, pond/pont, raad/raat, wand/want) occured in a plural condition (e.g., *raden en hennen*).

Fillers:

*troon en klok (throne and clock), *boon en klomp (bean and clog), *bed en gift (bed and gift), *blad en fout (leaveand error), *vod en scherm (rag and screen), *pad en schaal (path and scale), *brood en flank (bread and flank), *zaad en school (seed and school), *hoed en gracht (hat and canal), *kleed en schram (carpet and scratch), *naad en glas (seamand glass), *draad en frats (thread and face), kas en geit (cash register and goat), kies en gier (molar and vulture), *bijl en graan (axe and grain), *bel en graaf (bell and count), halm en leem (stalk and loam), hals en laan (neck and street), meel en juf (flour and teacher), meer en jurk (sea and dress), *bek en zool (mouth and sole), *bes en zoom (berry and hem), schim en nar (shadow and fool), schil en nek (bowl and nape)

All fillers in addition occured the two nouns in reversed order. The items marked with a * furthermore occured in a plural context (e.g., *naden en glasen*).

A9.2 Materials of the Second Perception Experiment (9b)

Assignment of experimental items to blocks:

blocks clitic 1 and final 1

itemnr	speaker
1,3,4,7,8,9	1,3
2,5,6,10,11,12	2,4
13	1,4

blocks clitic 2 and final 1

itemnr	speaker
1,3,4,7,8,9	2,4
2,5,6,10,11,12	1,3
13	2,3

block plural 1

itemnr	speaker
4,9	1,3
8,11	2,4

block plural 2

itemnr	speaker
4,9	2,4
8,11	1,3

A9.3 Results for the Voiced Responses ('d' = (4)+(5))

Voiced Responses in the three context conditions:

context	freq.	prop.
clitic:	655	.39
final:	656	.39
plural:	258	.50

Effect of context: $F(2;30) = 8.96$, $MSE = .11$, $p < .01$
This effect arose because subjects assigned almost no undecided responses in the plural condition, but only in the other contexts. A Tukey test for paired comparisons showed that the clitic and the final context condition did not differ from each other.

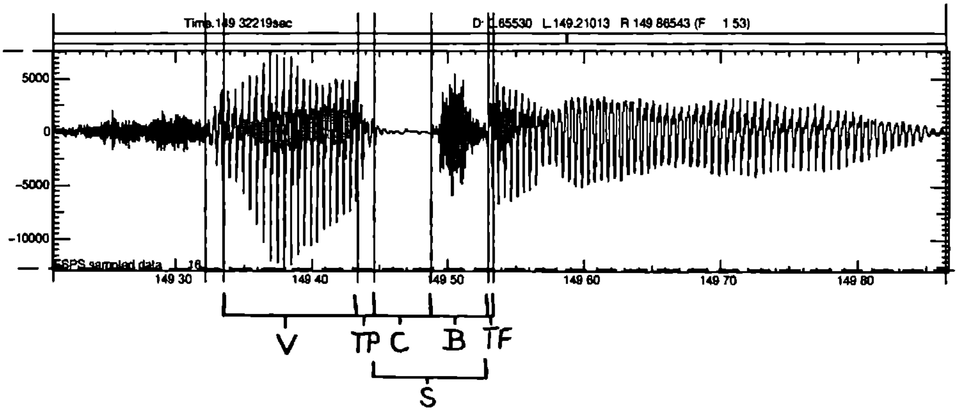
Effect of voicing: $F(1;15) = 563.47$, $MSE = .02$, $p < .01$
Interaction voicing by context: $F(2;30) = 317.14$, $MSE = .03$, $p < .01$
Effect of voicing in the plural condition $F(1;15) = 827.25$, $MSE = .03$, $p < .01$.

Effect of speaker: $F(3;45) = 17.53$, $MSE = .05$, $p < .01$.

Interaction speaker by context: $F(6.90) = 10.03$, $MSE = .04$, $p < .01$
Context by voicing within speaker 4: $F(2;30) = 109.81$, $MSE = .02$, $p < .01$.

A9.4 Example for a Segmented Speechwave

Item *graad en hol* (degree and cave) spoken by a female speaker



V = preceding vowel, TP = transition preceding the stop, C = stop closure, B = burst release, S = stop consonant, TF = transition following the stop

APPENDIX B: TESTS OF SIMPLE EFFECTS

B1 Experiment 1

Effect of relatedness at each SOA:
SOA 0 ms: $F_1(1;48) = 28.15$, $MSE = 5364$, $p < .01$; $F_2(1;9) = 16.4$, $MSE = 3672$, $p < .01$
SOA 150 ms: $F_1(1;48) = 75.12$, $MSE = 5364$, $p < .01$; $F_2(1;9) = 32.29$, $MSE = 5009$, $p < .01$

B2 Experiment 2

Interaction of relatedness and verb form for short and long IS at each SOA:
SOA 0 ms:
short IS: $F_1(1;48) = 2.02$, $MSE = 1186$, $p <.16$; $F_2(1;19) = 2.69$, $MSE = 710$, $p <.12$
long IS: $F_1(1;48) = 2.59$, $MSE = 1197$, $p <.11$; $F_2(1;19) = 1.75$, $MSE = 1389$, $p <.2$
SOA 150 ms:
short IS: $F_1(1;48) < 1$; $F_2(1;19) < 1$
long IS: $F_1(1;48) < 1$; $F_2(1;19) < 1$

Interaction of relatedness and length for infinitive and past tense targets at each SOA:
SOA 0 ms:
infinitive: $F_1(1;48) <1$; $F_2(1;19) < 1$
past tense: $F_1(1;48) = 11.34$, $MSE = 1034$, $p <.01$;
$F_2(1;19) = 7.71$, $MSE = 1211$, $p <.02$
SOA 150 ms:
infinitive: $F_1(1;48) <1$; $F_2(1;19) = 1.08$, $MSE = 887$, $p <.31$
past tense: $F_1(1;48) = 2.05$, $MSE = 1038$, $p <.16$;
$F_2(1;19) = 1.74$, $MSE = 982$, $p <.21$

B3 Experiment 3

Effect of relatedness at each SOA:
SOA -150 ms: $F_1(1;96) = 95.19$, $MSE = 9313$, $p < .01$; $F_2(1;19)=3.32$, $MSE = 1151$, $p <.09$
SOA 0 ms: $F_1(1;96) = 142.67$, $MSE = 9313$, $p < .01$; $F_2(1;19)= 12.06$, $MSE = 1725$, $p < .01$
SOA 150 ms: $F_1(1;96) = 145.45$, $MSE = 9313$, $p < .01$; $F_2(1;19)= 31.77$, $MSE = 2177$, $p < .01$
SOA 300 ms: $F_1(1;96) = 59.97$, $MSE = 9313$, $p < .01$; $F_2(1;19)= 6.79$, $MSE = 2094$, $p < .01$

Effect of length at each SOA:
SOA -150 ms: $F_1(1;96) = 14.08$, $MSE = 941$, $p < .01$; $F_2(1;19) = 8.84$, $MSE = 1197$, $p < .01$
SOA 0 ms: $F_1(1;96) = 15.6$, $MSE = 941$, $p < .01$; $F_2(1;19) = 10.97$, $MSE = 1089$, $p < .01$
SOA 150 ms: $F_1(1;96) = 22.37$, $MSE = 941$, $p < .01$; $F_2(1;19) = 25.0$, $MSE = 678$, $p < .01$
SOA 300 ms: all $F < 1$

Interaction verb form by relatedness by length at each SOA:
SOA -150 ms: all $F < 1$
SOA 0 ms: $F_1(1;96)=2.75$, $MSE = 958$, $p < .1$; $F_2(1;19) = 3.77$, $MSE = 552$, $p < .07$
SOA 150 ms: $F_1(1;96)=1.27$, $MSE = 958$, $p < .26$; $F_2(1;19) = 1.01$, $MSE = 954$, $p < .33$
SOA 300 ms: all $F < 1$

B4 Experiment 4

Effect of relatedness at each SOA:
SOA 0 ms: $F_1(1;72) = 42.3$, $MSE = 5092$, $p < .01$; $F_2(1;15) = 31.3$, $MSE = 4413$, $p < .01$
SOA 150 ms: $F_1(1;72) = 83.82$, $MSE = 5092$, $p < .01$; $F_2(1;15) = 61.01$, $MSE = 4492$, $p < .01$
SOA 300 ms: $F_1(1;72) = 9.67$, $MSE = 5092$, $p < .01$; $F_2(1;15) = 8.03$, $MSE = 3954$, $p < .01$

B5 Experiment 5a,b

Experiment 5a

Effect of relatedness at each SOA:
SOA 0 ms: $F_1(1;72) = 61.38$, $MSE = 4646$, $p < .01$;
$F_2(1;19) = 59.00$, $MSE = 3854$, $p < .01$
SOA 150 ms: $F_1(1;72) = 97.81$, $MSE = 4646$, $p < .01$;
$F_2(1;19) = 102.23$, $MSE = 3559$, $p < .01$
SOA 300 ms: $F_1(1;72) = 50.96$, $MSE = 4646$, $p < .01$;
$F_2(1;19) = 50.8$, $MSE = 3725$, $p < .01$

Effect of length at each SOA:
SOA -150 ms: $F_1(1;72) = 24.57$, $MSE = 1363$, $p < .01$; $F_2(1;19) = 18.99$, $MSE = 1417$, $p < .01$
SOA 0 ms: $F_1(1;72) = 14.94$, $MSE = 1363$, $p < .01$; $F_2(1;19) = 10.77$, $MSE = 1506$, $p < .01$
SOA 150 ms: $F_1(1;72) = 2.03$, $MSE = 1363$, $p < .16$; $F_2(1;19) = 1.68$, $MSE = 1332$, $p < .22$

Experiment 5b

Effect of relatedness at each SOA:
SOA -150 ms: $F_1(1;48) = 49.42$, $MSE = 3987$, $p < .01$; $F_2(1;11) = 39.53$, $MSE = 2382$, $p < .01$
SOA 0 ms: $F_1(1;48) = 87.96$, $MSE = 3987$, $p < .01$; $F_2(1;11) = 93.62$, $MSE = 1800$, $p < .01$

B6 Experiment 6

Effect in the Analysis over Items with Morphology as a Between-Items Factor:
Effect of relatedness: $F_2(1;26) = 111.22$, $MSE = 12229$, $p < .01$
Effect of SOA: $F_2(2;52) = 8.62$, $MSE = 4016$, $p < .01$
Interaction relatedness by length: $F_2(1;26) = 7.53$, $MSE = 2008$, $p < .02$
Interaction of relatedness by length at SOA 150 ms:
$F_2(1;26) = 5.69$, $MSE = 1750$, $p < .03$

Interaction of relatedness by length within simple targets:
$F_2(1;26) = 5.33$, $MSE = 2008$, $p < .03$

B7 Experiment 7

Effect of relatedness at each SOA:
SOA 0 ms: $F_1(1;48) = 49.42$, $MSE = 3987$, $p < .01$; $F_2(1;11) = 39.53$, $MSE = 2382$, $p < .01$
SOA 150 ms: $F_1(1;48) = 87.96$, $MSE = 3987$, $p < .01$; $F_2(1;11) = 93.62$, $MSE = 1800$, $p < .01$

B8 Reanalysis over Difference Scores from Pink Noise

Experiment 5a

Effect of length of related IS at each SOA:

SOA -150 ms:	$F_1(1;72) = 19.95$, $MSE = 1105$, $p < .01$;
	$F_2(1;19) = 15.97$, $MSE = 1107$; $p < .01$
SOA 0 ms:	$F_1(1;72) = 5.84$, $MSE = 1105$, $p < .02$;
	$F_2(1;19) = 4.61$, $MSE = 1138$; $p < .05$
SOA 150 ms:	all $F < 1$

Effect of length of unrelated IS at each SOA:
SOA -150 ms: $F_1(1;72) = 7.86$, $MSE = 1556$, $p < .01$; $F_2(1;19) = 5.47$, $MSE = 1798$, $p < .03$
SOA 0 ms: $F_1(1;72) = 9.46$, $MSE = 1556$, $p < .01$; $F_2(1;19) = 9.05$, $MSE = 1311$, $p < .01$
SOA 150 ms: $F_1(1;72) = 2.28$, $MSE = 1556$, $p < .14$; $F_2(1;19) = 2.19$, $MSE = 1289$, $p < .16$

B9 Acoustic Measurements

Effect of voicing within final and clitic context:

Preceding vowel:	clitic: $F(1;12) = 2.02$, $MSE = 18069$, $p < .19$;
	final: $F < 1$
Stop Closure:	clitic: $F < 1$;
	final: $F(1;12) = 6.80$, $MSE = 28271$, $p < .03$
Burst:	clitic: $F(1;12) = 1.39$, $MSE = 6010$, $p < .27$;
	final: $F < 1$
Stop Consonant:	clitic: $F < 1$;
	final: $F(1;12) = 1.00$, $MSE = 74446$, $p < .34$

SAMENVATTING

Het onderwerp van dit proefschrift is de rol van de syllabe in de produktie van gesproken taal. Voor de meeste mensen zijn syllaben intuïtief plausibele elementen. In een woord als "kanarievogel" zal iedereen zonder veel moeite de syllaben ka-na-rie-vo-gel herkennen. In gesproken taal gebeurt het vaak dat de grenzen tussen de verschillende woorden en syllaben verschoven zijn, zoals in het geval van zogenaamd enclitische vormen, waar een voornaamwoord of een ander zwak funktiewoord versmolten is met bijvoorbeeld een werkwoord: De zin "Ik kook het" wordt doorgaans uitgesproken als "Ik kook 't". Dat wil zeggen, *kook het* valt uiteen in de syllaben $(ko:)_\sigma(kət)_\sigma$. Dit is beslist geen ongebruikelijk verschijnsel, aangezien talen de neiging hebben om hun syllaben niet met een klinker te laten beginnen. De laatste medeklinker van een woord komt doorgaans terecht aan het begin van de eerste syllabe van het daaropvolgende woord.

De meeste onderzoekers gaan er van uit dat syllaben niet als kant en klare eenheden uit het mentale lexicon worden opgehaald, maar tijdens het fonologisch encoderingsproces geconstrueerd worden. Maar hoe een spreker tijdens het spreken nu precies een syllabestructuur opbouwt is nog een open vraag.

Levelt's model van fonologische encodering tijdens het spreken is het enige model in zijn soort dat verder gaat dan de produktie van het enkele woord, en dat ook daadwerkelijk de produktie van aaneengesloten woorden behandelt (Levelt, 1989, 1992). Volgens dit model genereert de zogenaamde segmentele spellout procedure eerst een geordende reeks van fonologische segmenten die in de beoogde uiting zullen voorkomen. In het voorbeeld *kook het* zijn dit de segmenten /k, o:, k, ə, t/. Vervolgens worden deze segmenten gekoppeld aan een langs onafhankelijke weg gegenereerde prosodische structuur. Deze structuur, die aangeeft hoeveel syllaben er zijn en waar de klemtoon ligt, kan meerdere woorden omvatten. In het geval van *kook het* is er slechts één prosodische structuur die bestaat uit een beklemtoonde syllabe gevolgd door een onbeklemtoonde. De segmenten worden aan deze prosodische structuur verbonden volgens bepaalde algemene principes, zoals bijvoorbeeld het 'Maximal Onset' principe. Dit laatste principe zegt dat medeklinkers een voorkeur hebben voor het begin van een syllabe, waarbij een syllabestructuur wordt aangenomen waarin de "onset" verwijst naar die medeklinkers die in een syllabe voorafgaan aan de klinker, en waarin de "coda"-positie gedefinieerd wordt als die medeklinkers die na deze klinker komen. Volgens dit principe zal de /k/ aan het eind van het woord *kook* terecht komen op de onset-positie van de tweede syllabe, zoals in $(ko:)_\sigma(kət)_\sigma$.

Op basis van het model van Levelt doe ik twee voorspellingen over de wijze waarop sprekers syllaben genereren. Deze twee voorspellingen worden in de twee afzonderlijke delen van dit proefschrift getoetst.

De eerste voorspelling betreft het tijdsverloop van het fonologisch encoderen. Volgens Levelt's model worden syllaben pas op een laat moment in dit proces gegenereerd, d.w.z. pas nadat de fonologische segmenten alsmede de prosodische structuur beschikbaar zijn gekomen. Op een vroeger moment zouden wel de segmenten beschikbaar moeten zijn, maar zou de syllabe als eenheid nog geen rol mogen spelen.

De tweede voorspelling betreft het aantal nivo's van syllabestructuur dat bij het fonologisch encoderen betrokken is. Men zou kunnen aannemen dat een spreker eerst de syllaben van de te produceren individuele woorden construeert, bijvoorbeeld *kook* $(\text{ko:k})_\sigma$ en *het* $(\text{ət})_\sigma$. Vervolgens zouden deze syllaben dan achter elkaar worden gezet, waarbij de medeklinkers die voorafgaan aan een syllabe die met een klinker begint op de onset-positie van die syllabe terecht komen, conform het eerder genoemde Maximal Onset principe. Het resultaat is een enclitische vorm, in dit geval $(\text{ko:})_\sigma(\text{kət})_\sigma$. Dit type proces bevat twee nivo's van syllabestructuur: De 'diepe' en abstracte lexicale syllaben $(\text{ko:k})_\sigma$ en $(\text{ət})_\sigma$, en de uiteindelijke 'oppervlakte' syllaben $(\text{ko:})_\sigma$ en $(\text{kət})_\sigma$. In Levelt's model van fonologisch encoderen, daarentegen, wordt geen tussenliggend nivo van lexicale syllaben verondersteld. In plaats daarvan zal de spreker direct de oppervlakte syllaben genereren. Gezien het tempo waarmee wij syllaben produceren - zo'n 3 tot 5 syllaben per seconde - lijkt het plausibel om niet uit te gaan van de constructie van een tussenliggend abstract nivo van lexicale syllaben die nooit in de gesproken taal terecht komen.

Het eerste gedeelte van het proefschrift houdt zich bezig met het tijdsverloop van syllabe-constructie, en toetst de voorspelling dat syllaben pas op een laat moment tijdens het fonologisch encoderen beschikbaar zijn. In verschillende experimenten werd eerst aan deelnemers gevraagd om afzonderlijke paren van zelfstandig naamwoorden en semantisch gerelateerde werkwoorden uit het hoofd te leren, bijvoorbeeld *soep - koken*. Vervolgens was het hun taak om deze werkwoorden uit te spreken zodra het gerelateerde zelfstandig naamwoord op een computerscherm verscheen. Zodra het woord *soep* op het scherm verscheen moesten ze bijvoorbeeld *koken* zeggen. De werkwoorden die deelnemers moesten uitspreken zal ik in het vervolg 'doelwoorden' noemen. In sommige experiment-blokken werd aan de deelnemers gevraagd om de verledentijdsvorm van het werkwoord te benoemen, zoals *kookte*, in andere delen de enclitische vorm, zoals *kook het*. Deze doelvormen hadden dezelfde initiële segmenten, maar ze verschilden in hun eerste syllabe, zoals bij *kookte* $(\text{ko:k})_\sigma(\text{tə})_\sigma$ vergeleken met *koken* $(\text{ko:})_\sigma$ $(\text{kən})_\sigma$ of *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$. Deelnemers kregen, terwijl ze een doelwoord aan het benoemen waren, steeds ook een auditieve syllabestimulus te horen. Deze stimulus kon fonologisch gerelateerd zijn aan het doelwoord, zoals /ko:/ of /ko:k/, maar hij kon ook ongerelateerd zijn aan dat doelwoord, d.w.z. andere segmenten bevatten, zoals /le:/ of /le:n/. Eerder onderzoek had al laten zien dat deelnemers sneller een woord konden benoemen wanneer dat woord vergezeld werd door een fonologisch gerelateerde auditieve stimulus dan wanneer het vergezeld werd door een ongerelateerde stimulus. We verwachtten dit effect daarom ook in onze eigen experimenten. Verder kon een stimulus kort zijn, zoals /ko:/, of lang, zoals /ko:k/. De korte gerelateerde stimuli kwamen overeen met de eerste syllabe in de infinitief- en de enclitische vormen (bijvoorbeeld /ko:/ in *koken* of *kook het*), terwijl de lange gerelateerde stimuli overeen kwamen met de eerste syllabe van de verledentijdsvormen (bijvoorbeeld /ko:k/ in *kookte*).

Relatief ten opzichte van de visuele presentatie van het zelfstandig naamwoord (bijvoorbeeld *soep*) dat de benoeming van het doelwoord (bijvoorbeeld *koken*) eliciteerde, kon de auditieve stimulus op verschillende momenten worden aangeboden, namelijk voor, tegelijkertijd met, of na dit naamwoord. Deze experimentele manipulatie is van cruciaal belang in relatie tot Levelt's model van fonologische encodering, aangezien hierin verondersteld wordt dat de spreker in de vroege stadia van spraakproduktie al wel segmenten, maar nog geen syllaben gegenereerd heeft. Om deze reden zou een vroeg aangeboden auditieve stimulus dan ook geen syllabe-afhankelijke invloed mogen hebben op de fonologische encodering van het doelwoord. Deze vroege stimulus zou daarentegen wel een segment-afhankelijke invloed moeten hebben op dit proces. In dit verband kunnen we de volgende voorspelling afleiden: Een vroeg aangeboden, fonologisch gerelateerde auditieve stimulus zal, naarmate hij meer segmenten deelt met het te produceren doelwoord, in sterkere mate de tijd bekorten die nodig is om dat doelwoord daadwerkelijk te benoemen, dit in vergelijking tot een ongerelateerde auditieve stimulus van dezelfde lengte. Met andere woorden, deelnemers zouden meer baat moeten hebben bij gerelateerde stimuli die lang zijn dan bij gerelateerde stimuli die kort zijn. Dit zou moeten gelden voor alle doelwoorden, ongeacht hun syllabestructuur. Dit voorspelde effect noem ik het 'segment-overlappingseffect'.

Gegeven het tijdsverloop van fonologische encodering in Levelt's model kunnen we een tweede voorspelling doen: Een later gepresenteerde auditieve stimulus zou wel een syllabe-afhankelijke invloed moeten hebben op de benoeming van het doelwoord, aangezien syllaben pas later in dat benoemingsproces beschikbaar komen. Voor dit latere tijdstip voorspelt het model dus een 'syllabe-identiteitseffect'. Auditieve gerelateerde stimuli die overeen komen met de eerste syllabe van een doelwoord zouden het benoemingsproces meer moeten versnellen dan gerelateerde stimuli die niet met die eerste syllabe overeen komen, opnieuw in vergelijking tot een ongerelateerde stimulus van dezelfde lengte. Met andere woorden, wanneer de deelnemers infinitiefvormen of enclitische vormen produceren, en dit zijn vormen die alle beginnen met een korte syllabe, zoals $(\text{ko:})_\sigma(\text{kən})_\sigma$ of $(\text{ko:})_\sigma(\text{kət})_\sigma$, dan zouden ze meer baat moeten hebben bij een gerelateerde stimulus die kort is dan bij een die lang is. Wanneer ze daarentegen verledentijdsvormen produceren, en dit zijn vormen die alle beginnen met een lange syllabe, zoals $(\text{ko:k})_\sigma(\text{tə})_\sigma$, dan zouden ze meer baat moeten hebben bij een lange dan bij een korte gerelateerde stimulus.

Zoals verwacht benoemden de deelnemers de doelwoorden sneller wanneer de auditieve stimuli fonologisch aan deze woorden gerelateerd waren dan wanneer ze ongerelateerd waren. Daarnaast waren de deelnemers in het algemeen langzamer bij een lange dan bij een korte auditieve stimulus, ongeacht of deze nu gerelateerd was of niet. Dit laat zien dat de experimentele manipulaties daadwerkelijk op het proces van fonologische encodering ingrepen, aangezien de deelnemers bijvoorbeeld gevoelig bleken te zijn voor het verschil tussen korte en lange stimuli als /ko:/ en /ko:k/, een verschil van slechts een segment. Bovendien is met twee kontrole-experimenten aangetoond dat de resultaten niet door morfologische of lexicale variabelen kunnen zijn veroorzaakt.

Maar hoe zat het nu met de rol van de syllabe? Een van deze experimenten, Experiment 2, leverde een resultaat op dat als evidentie voor de voorspellingen van Levelt's model van fonologische encodering zou kunnen worden opgevat. Dit experiment onderzocht de produktie van infinitief- en verledentijdsvormen, zoals *koken* en *kookte*. Het bleek dat korte en lange gerelateerde stimuli de benoeming van infinitiefvormen in gelijke mate versnelden, terwijl de benoeming van verledentijdsvormen meer baat had bij lange dan bij korte gerelateerde stimuli. Een dergelijk patroon kan worden voorspeld zodra we aannemen dat het 'segment-overlappingseffect' en het 'syllabe-identiteitseffect' niet op verschillende tijdstippen optreden, maar tegelijkertijd. Volgens het eerste effect zouden lange stimuli effectiever moeten zijn dan korte stimuli bij de benoeming van beide werkwoordvormen. Volgens het tweede effect zouden korte stimuli effectiever moeten zijn bij de benoeming van infinitiefvormen, terwijl de lange stimuli effectiever zouden moeten zijn bij de benoeming van verledentijdsvormen. Als beide effecten additief en ongeveer even groot zijn, dan zouden ze elkaar moeten opheffen bij de benoeming van infinitiefvormen, terwijl ze bij de benoeming van verledentijdsvormen juist het verschil tussen korte en lange stimuli zouden moeten versterken. Dit is precies het patroon van de resultaten van Experiment 2. Anders dan verwacht werd echter bij een latere presentatie van de stimulus nog steeds fonologische facilitatie gevonden, terwijl het zojuist beschreven patroon verdwenen was. Volgens het model zou bovengenoemd patroon uitsluitend op een laat tijdstip moeten worden gevonden, wanneer de syllabe als eenheid beschikbaar komt.

Het zou natuurlijk kunnen dat sprekers werkwoorden zoals *koken* inmiddels zo vaak hebben uitgesproken dat het fonologisch proces gewoon te snel verloopt om nog effecten te kunnen vinden. Ik heb daarom geprobeerd het resultaat te repliceren in een experiment met enclitische doelvormen, vormen die waarschijnlijk minder vaak zijn gebruikt. Enclitische doelvormen met een korte eerste syllabe, zoals *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$, gedroegen zich als de infinitiefvormen in Experiment 2: de benoemingstijd werd in gelijke mate door korte en lange stimuli beïnvloed. Het bijbehorende experiment met doelvormen die begonnen met een lange syllabe, zoals *kookt het* $(\text{ko:k})_\sigma(\text{tət})_\sigma$ leverde echter een ander resultaat op dan wat in Experiment 2 bij de verledentijdsvormen was gevonden. Aangezien gerelateerde lange stimuli als /ko:k/ meer segmenten met de doelvormen gemeen hebben dan korte stimuli als /ko:/, en bovendien ook nog met de eerste syllabe van deze doelvormen overeenkomen, was de verwachting dat ze de benoeming veel meer zouden versnellen dan de gerelateerde korte stimuli. Het effect van korte en lange stimuli was echter gelijk. Het was overigens wel zo dat de deelnemers in het algemeen meer moeite hadden met het produceren van de *kookt het* vormen dan met het produceren van de *kook het* vormen, en dit kan het resultaat hebben beïnvloed.

Samengevat: het is nog steeds een open vraag of syllaben pas in een laat stadium van het fonologisch encoderingsproces gegenereerd worden. Het resultaat van Experiment 2 was echter veelbelovend. De mate waarin deelnemers baat hadden bij korte en lange gerelateerde stimuli hing daadwerkelijk af van de syllabestructuur van de doelwoorden. Dit wijst erop dat de syllabe een relevant element is in het proces van spraakproduktie. Het precieze tijdsverloop waarmee deze elementen beschikbaar komen zal in verder onderzoek nader moeten worden bestudeerd.

Het tweede deel van dit proefschrift gaat over de tweede voorspelling van Levelt's model van fonologische encodering, dat stelt dat er slechts één nivo van syllabestructuur is, namelijk die van oppervlakte syllaben. Deze voorspelling impliceert dat de spreker die een enclitische vorm zoals *kook het* $(\text{ko:})_\sigma(\text{kət})_\sigma$ produceert, nooit op een abstract tussenliggend nivo van representatie syllaben zal genereren die horen bij de afzonderlijke lexicale elementen. Dat betekent dat de /k/ in *kook het* nooit een coda-positie zal innemen, hetgeen wel degelijk het geval zou zijn bij het enkele woord *kook* $(\text{ko:k})_\sigma$. De spreker zal daarentegen direct de oppervlakte syllaben genereren, waarbij de /k/ terecht komt in de onset-positie van de volgende syllabe, zoals in $(\text{ko:})_\sigma(\text{kət})_\sigma$.

Om deze voorspelling te kunnen toetsen ben ik op zoek gegaan naar een verschijnsel dat van invloed is op medeklinkers in coda-positie. Medeklinkers die bij enclitische vormen in onset-positie terecht komen (zoals de tweede /k/ in *kook het*) zouden niet mogen worden beïnvloed door een wetmatigheid die alleen betrekking heeft op coda-posities. Wanneer dergelijke medeklinkers echter wel sporen van een coda-gebaseerde wetmatigheid zouden vertonen, dan zouden we moeten concluderen dat deze medeklinkers, alvorens ze op hun uiteindelijke onset-positie terecht kwamen, eerst een coda-positie hebben ingenomen. Dit zou evidentie zijn voor een tussenliggend nivo van syllabestructuur, een nivo dat voorafgaat aan dat van de oppervlakte syllaben.

'Syllable-final devoicing' is zo'n verschijnsel. Nederlandse syllaben kunnen met een stemhebbende of met een stemloze klank beginnen. Sommige woorden verschillen alleen in het stemhebbende danwel stemloze karakter van een medeklinker, zoals bij *raden* en *raten*. In de coda-positie van een Nederlandse syllabe kunnen echter uitsluitend stemloze plof- en wrijfklanken voorkomen. Dit heeft tot gevolg dat de enkelvoudsvormen van *raden* en *raten* wat betreft hun klankvorm identiek zijn: $(\text{ra:t})_\sigma$. Terwijl het stemcontrast duidelijk overeind blijft in de meervoudsvormen, en wordt opgeheven in de enkelvoudsvormen, zijn enclitische vormen relevant voor het tussenliggend nivo: Op grond van de voorspelling dat oppervlaktesyllaben de enige syllaben zijn die tijdens fonologische encodering gegenereerd worden, moet een plofklank die stemhebbend is in meervoudsvormen ook stemhebbend zijn in enclitische vormen, aangezien deze consonant nooit in een coda-positie terecht komt waar het stemloos zou kunnen worden. In de enclitische vorm *raad en* bijvoorbeeld, waar de /d/ in onset-positie van de tweede syllabe terecht komt, zou deze consonant stemhebbend moeten zijn: $(\text{ra:})_\sigma(\text{dən})_\sigma$. Anders dan in Levelt's model verondersteld wordt, zouden sprekers eerst de syllaben kunnen genereren die horen bij de afzonderlijke lexicale elementen, zoals *raad* en *en*. De plofklank aan het einde van de syllabe zou dan stemloos kunnen worden voordat deze consonant in een onset-positie van de oppervlakte syllabestructuur terecht komt. Dit proces, waarin de plofklank van positie in de syllabestructuur verandert, wordt 'hersyllabificatie' genoemd.

In twee perceptie-experimenten heb ik gebruik gemaakt van paren van zelfstandig naamwoorden die alleen verschilden in of hun laatste consonant stemhebbend of stemloos was (zoals *raad* en *raat*). Dit stemcontrast werd opgeheven indien de plofklank aan het einde van de uiting, en daarmee ook aan het einde van een syllabe stonden, zoals in *olm en raat* of *olm en raad*. In deze context konden de deelnemers geen onderscheid maken tussen woorden als *raad* en *raat*. In de meervoudsvormen, daarentegen, bleef het stemcontrast bestaan, aangezien de plofklanken hier in een onset-positie stonden. Deelnemers namen stemhebbende plofklanken waar als stemhebbend, en stemloze plofklanken als stemloos. Belangrijk was dat het stemcontrast in enclitische vormen niet te horen was, hoewel de plofklank hier in onset-positie terecht kwam. Net als in de gevallen waar de plofklank aan het einde van de uiting stond, konden de deelnemers geen onderscheid maken tussen *raad* en *raat*. Akoestische metingen op de stimuli die in deze perceptie-experimenten waren gebruikt lieten hetzelfde resultatenpatroon zien. In een theorie die twee nivo's van syllabestructuur veronderstelt kunnen deze resultaten verklaard worden door aan te nemen dat 'syllable-final devoicing', het proces dat er voor zorgt dat stemhebbende plofklanken in coda-positie stemloos worden, zijn werk doet op het nivo waar de syllaben nog met de afzonderlijke woorden overeenkomen, en dus voordat hersyllabificatie een coda consonant verplaatst naar de onset-positie van de volgende syllabe (indien deze met een klinker begint). In Levelt's model van fonologische encodering, daarentegen, is het niet eenvoudig te verklaren waarom plofklanken in onset-posities van enclitische vormen zich gedragen alsof ze in coda-positie staan, omdat in dit model immers alleen een nivo van oppervlakte syllaben wordt verondersteld. In de oppervlaktestructuur komt de plofklank, bijvoorbeeld de /d/ in *raad en*, voor in onset-positie. Op geen enkel moment tijdens het fonologisch encoderingsproces komt deze plofklank voor in coda-positie.

De resultaten van Deel 2 zijn relevant voor een aantal kwesties in de fonologie, met name met betrekking tot de vraag hoe sterk fonologische theorievorming zich op oppervlakte-structuur zou moeten oriënteren, en met betrekking tot de status van hersyllabificatie. Sterk oppervlakte-georiënteerde fonologische theorieën en theorieën die zonder hersyllabificatie werken hebben soortgelijke problemen als Levelt's model om de resultaten in Deel 2 eenvoudig te verklaren. Bovendien zijn de resultaten relevant voor de relatie tussen fonologie en fonetiek. Het gaat om de vraag of een contrast dat op fonologisch nivo opgeheven is ook op fonetisch nivo geen sporen achterlaat. Een voorbeeld voor zo'n contrast is 'syllable-final devoicing'. Op fonologisch nivo wordt het contrast tussen stemhebbende en stemloze plof- of wrijfklanken in coda-positie opgeheven: Deze klanken zijn in coda-positie altijd stemloos. Sommige onderzoekers beweren dat op fonetisch nivo toch een verschil blijft bestaan tussen klanken die oorspronkelijk stemhebbend waren en klanken die steeds stemloos waren. De akoestische metingen in Deel 2 lieten zien dat dit niet altijd het geval hoeft te zijn.

Wat de fonologische encodering tijdens spraakproduktie betreft laten de in Deel 1 beschreven experimenten zien dat ook auditieve syllabestimuli een invloed hebben op het fonologische proces. Eerder soortgelijk onderzoek met auditieve stimuli had tot nu toe uitsluitend hele woorden gebruikt. Ook bleek de wijze waarop benoemingen geéliciteerd werden (m.b.v. een semantischeassociaten leertaak, zie Meyer, 1990, 1991) goed te combineren met dergelijke auditieve stimuli. Dit voegt een nieuwe variant toe aan het beperkte repertoire van onderzoeksmethodes waarmee spraakproduktie experimenteel kan worden onderzocht.

Ter verklaring van de bevindingen in Deel 2 bespreek ik een aantal verschillende manieren waarop de experimentele resultaten met het fonologisch encoderingsmodel van Levelt kunnen worden verenigd. Allereerst zou men kunnen stellen dat "syllable-final devoicing" niet aangrijpt op de laatste positie van een syllabe, maar op dat van een woord. Dit kan verklaren waarom de /d/ in *raad en* stemloos wordt. Ten tweede is het mogelijk te veronderstellen dat alle morfologisch complexe woorden, bijvoorbeeld ook meervoudsvormen en verledentijdsvormen, in het lexicon zijn opgeslagen, tezamen met de informatie over hun stemeigenschappen. Als alle vormen eenvoudigweg in het lexicon zijn opgeslagen, dan hoeft er tijdens het spraakproduktieproces geen stemverlies meer plaats te vinden. Een derde mogelijkheid is om te veronderstellen dat stemverlies tijdens fonologische encodering zowel optreedt bij het einde van syllaben alsook bij dat van woorden. Mijn eigen voorkeur gaat uit naar een vierde mogelijkheid, namelijk een waarin tijdens fonologische encodering toch hersyllabificatie optreedt. Dit zou betekenen dat er meerdere nivo's van syllabestructuur bestaan. Een consonant die op het nivo van lexicale syllabestructuur in coda-positie staat, en die daar gevolgd wordt door een syllabe die begint met een klinker, zal hersyllabificeren naar de onset-positie van die volgende syllabe. Verdere evidentie hiervoor treffen we aan in andere talen, zoals Spaanse dialecten of het Frans dat gesproken wordt in Québéc, talen die eveneens regelmatigheden vertonen die op een coda-element ingrijpen voordat dit element op de volgende onset-positie terecht komt. Ook kan hersyllabificatie in het model van fonologische encodering opgenomen worden zonder dat dit veel extra mentale belasting met zich meebrengt: Het segment waar het om gaat bevindt zich aan de rand van een lexicaal element, en hoeft alleen maar van de coda-positie van de ene syllabe naar de onset-positie van een volgende syllabe verplaatst te worden. Als we hersyllabificatie in het model toelaten kunnen we de resultaten voor de enclitische vormen verklaren.

Der Rheinländer denkt, wie er spricht,
nein, spricht wie er denkt.

(Beikircher, Himmel un Ääd)

Curriculum Vitae

Monika Baumann was born in Aachen on March, 27, 1967, where in Summer 1986 she graduated from the "St Ursula Gymnasium".
In October 1986 she went to Cologne to study German Language and Literature, General Linguistics, and Political Science. For three years, she worked as a student assistant at the German Department. In Summer 1992, she finished her study "with distinction" as a Master of Arts. Her M.A. Thesis was a theoretical and experimental study on pauses and lengthening phenomena in speech production. The experiment reported in the M.A. Thesis had been planned and prepared during an internship of five months at the Max Planck Institute for Psycholinguistics, Nijmegen, in 1991.
In September 1992 she received a three-years grant of the Max Planck Society and started as a PhD-student at the Max Planck Institute for Psycholinguistics. The research reported in this Doctoral Thesis results from these three years.
Since October 1995, Monika Baumann is a scientific staff member in a project on developmental dyslexia at the University of Hamburg. The project is funded by the German Research Foundation.

ISBN 90 9008946 2